

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US2005/007410

International filing date: 07 March 2005 (07.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/550,268
Filing date: 05 March 2004 (05.03.2004)

Date of receipt at the International Bureau: 27 July 2006 (27.07.2006)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1493637

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

July 18, 2006

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/550,268

FILING DATE: *March 05, 2004*

RELATED PCT APPLICATION NUMBER: *PCT/US05/07410*

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US60/550,268*



Certified by

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

PATENT
ARCD:405USP1

PROVISIONAL
APPLICATION FOR UNITED STATES LETTERS PATENT
for
METHODS AND COMPOSITIONS RELATING TO THE
PHARMACOGENETICS OF ABCC2 GENE VARIANTS
by
Mark J. Ratain
Federico Innocenti
Deanna L. Kroetz
and
Samir Undevia

CERTIFICATE OF EXPRESS MAIL	
Express Mail No.:	EV 414834592 US
Date of Deposit:	March 5, 2004

BACKGROUND OF THE INVENTION

The government may own rights in the present invention pursuant to grant number GM61393 from the National Institutes of Health.

1. Field of the Invention

The present invention relates generally to the fields of molecular genetics, pharmacogenetics, and cancer therapy. In particular, the present invention is directed to methods and compositions for detecting polymorphisms and correlating the presence or absence of certain polymorphisms with toxic effects of chemotherapies. More specifically, the present invention is directed to methods and compositions for determining the presence or absence of polymorphisms within an ABCC2 gene and correlating these polymorphisms with toxic effects of ABCC2 substrates, as well as evaluating the risk of an individual for developing toxicity to an ABCC2 substrate. In some embodiments, the invention concerns methods and compositions for predicting or anticipating the level of toxicity caused by an ABCC2 substrate, such as irinotecan, in a patient. Such methods and compositions can be used to evaluate whether irinotecan-based therapy, or therapy involving other ABCC2 substrates, may pose toxicity problems if given to a particular patient. Alterations in suggested therapy may ensue if a toxicity risk is assessed.

2. Description of Related Art

ATP-binding cassette (ABC) genes represent the largest family of transmembrane proteins that bind ATP and use the energy to drive the transport of various molecules across cell membranes. The products of the ABC genes are known to influence oral absorption and disposition of a wide variety of drugs and play a role in the resistance of malignant cells to anticancer agents (Sparreboom *et al.*, 2000).

ABCC2, a member of the ABC gene family, functions as the major exporter of organic anions from the liver into the bile. In addition, ABCC2 is expressed on the apical membrane of epithelial cells such as enterocytes, renal proximal tubule epithelia, and gall bladder epithelia. ABCC2 is also expressed in some tumor tissues such as ovarian carcinoma, colorectal carcinoma, leukemia, mesothelioma, and hepatocarcinoma; and it

has been suggested that tumor cells overexpressing ABCC2 acquire multidrug resistance (MDR) (Borst *et al.* (1999); Borst *et al.* (2000)).

ABCC2 substrates include intracellularly formed glucuronide and reduced glutathione (GSH)—conjugates of clinically important drugs (Suzuki *et al.*, 1998). In
5 addition, ABCC2 is also involved in the biliary excretion of non-conjugated anionic drugs such as irinotecan (CPT-11).

Irinotecan is an antineoplastic drug used in the treatment of colon cancer. Irinotecan hydrolysis by carboxylesterase-2 (CES-2) is responsible for its activation to SN-38 (7-ethyl-10-hydroxycamptothecin), a topoisomerase I inhibitor of much higher
10 potency than irinotecan. The main inactivating pathway of irinotecan is the biotransformation of active SN-38 into inactive SN-38 glucuronide (SN-38G) by UDP-glucuronosyltransferase 1A1 (UGT1A1) (Iyer *et al.*, 1998).

Despite its efficacy in treating metastatic colon cancer and its broad spectrum of activity in other tumor types, irinotecan treatment is associated with significant toxicity.
15 The main severe toxicities of irinotecan are delayed diarrhea and myelosuppression. In the early single agent trials, grade 3-4 diarrhea occurred in about one third of patients and was dose limiting (Negoro *et al.*, 1991; Rothenberg *et al.*, 1993). Its frequency varies from study to study and is also schedule dependent. The frequency of grade 3-4 diarrhea in the three-weekly regimen (19%) is significantly lower compared to the weekly
20 schedule (36%, Fuchs *et al.*, 2003). In addition to diarrhea, grade 3-4 neutropenia is also a common adverse event, with about 30-40% of the patients experiencing it in both weekly and three-weekly regimens (Fuchs *et al.*, 2003; Vanhoefer *et al.*, 2001). Fatal events during irinotecan treatment have been reported. A high mortality rate of 5.3% and 1.6% was reported in the weekly and three-weekly single agent irinotecan regimens,
25 respectively (Fuchs *et al.*, 2003).

Interpatient differences in systemic formation of SN-38G have been shown to have clear clinical consequences in patients treated with irinotecan. Patients with higher glucuronidation of SN-38 are more likely to be protected from the dose limiting toxicity of diarrhea in the weekly schedule (Gupta *et al.*, 1994).

30 Improved methods and compositions for the evaluation of risk for irinotecan toxicity in an individual are still needed. Clearance of irinotecan and its metabolites by

ABCC2 represents a mechanism to protect patients from the toxic effects. However, the problem of identifying the effects of various polymorphisms on drug clearance by ABCC2 remains. Resolving these problems would provide novel methods and compositions for the evaluation of risk for toxicity to irinotecan as well as for numerous other drugs that are substrates for ABCC2.

SUMMARY OF THE INVENTION

The present invention is based on identification and characterization of correlations between genotype of the *ABCC2* gene and phenotype relating to the activity of ABCC2. Thus, the present invention provides methods and compositions that exploit correlations between genotype and phenotype concerning ABCC2. It is contemplated that such methods and compositions have diagnostic, prognostic, and therapeutic applications.

The present invention involves methods for determining the level of ABCC2 activity in a patient. This method can be used to predict what the level of ABCC2 activity is in a patient based on genotypic analysis. In some embodiments, the method involves determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 on one or both alleles is indicative of a normal level of ABCC2 activity.

Additional methods of the invention include a method for predicting tumor response to an anticancer agent that is an ABCC2 substrate in a cancer patient comprising determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 on one or both alleles is indicative of a greater chance of a reduced antitumor response to the anticancer agent. The probability of a reduced antitumor response is increased with respect to persons who do not have a C at position 3972. The determination of a T on both alleles at position 3972 in the *ABCC2* gene is indicative of a greater chance of an antitumor response or of a better antitumor response than would be expected as compared to a person with a C at position 3972.

The term "antitumor response" means a response that results in a favorable therapeutic outcome with respect to a tumor. Examples of such an outcome include, but are not limited to, reduction in tumor size, retardation of tumor growth or proliferation, inhibition of metastasis, reduction in number of metastasis, inhibition of tumor

vasculature, inhibition of tumor growth rate, promotion of apoptosis of tumor cells, induction of tumor cell death or killing, promotion of remission of cancer growth, and extended survival. Thus, a reduced antitumor response means the patient may exhibit no response to the drug or that the response is less favorable than would be expected for someone with a TT genotype at position 3972. It will understood that the prediction of a reduced antitumor response may lead to an increased dosage (increased concentration, increased administration and/or both) and/or more aggressive treatment regimen than would have been the case for someone with the TT genotype. This altered treatment may overcome the predicted reduced antitumor response. Thus, embodiments of the invention further include adjusting dosage (concentration and/or administration (timing and/or frequency)) or route of administration of the anticancer agent or altering the treatment regimen overall. In some cases, the time between treatment regimens may be altered. In specific embodiments, the anticancer agent is irinotecan.

Other methods of the invention concern a method for determining dosage of an ABCC2 substrate for a patient comprising: a) determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 on one or both alleles indicates a higher dosage of the substrate than is indicated for a patient with a T at position 3972 in both alleles of the *ABCC2* gene. In some embodiments, the ABCC2 substrate is selected from the group of substrates consisting of cysteinyl leukotrienes, glutathione and glutathione conjugates, glucuronide conjugates, sulfated conjugates, bile salt conjugates, bromosulfophthalein, and dibromosulfophthalein (see Table 1). Identified in Table 1 are substrates that are administered as drugs to patients. Determining the dosage of any of these drugs is specifically contemplated as part of the invention. In some cases, the dosage that would be given to a patient is modified based on the genotyping results based on methods of the invention. In certain embodiments, the substrate is irinotecan, SN-38, APC, and/or SN-38G. Methods of the invention also include prescribing a dosage of the anticancer agent, such as irinotecan, based on the determination of the sequence at position 3972 in one or both alleles of the *ABCC2* gene. It is contemplated that a patient is given a different dosage than he or she would have otherwise received had the genotyping not been performed. Thus, in some embodiments

of the invention, a typical dosage is adjusted for a particular person (individualized therapy).

Methods of the invention also include monitoring for toxicity or adverse events once the ABCC2 substrate is administered, and possibly, adjusting or modifying dosage based on the those results. Toxicity indicators or indicators of adverse events include diarrhea, neutropenic fever, other hematologic toxicities, as well as known non-hematologic toxicities.

The present invention also concerns a method for predicting a clearance rate for irinotecan in a patient. The method involves determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 in one or both alleles is indicative of a normal clearance rate for irinotecan. Again, "normal" is with respect to the level of clearance that is expected for persons with the TT haplotype at position 3972. In additional embodiments, the clearance rate is determined empirically in that patient based on techniques that are well known to those of skill in the art. Identification of a T at position 3972 on both alleles of the *ABCC2* gene is indicative of a lower than normal clearance rate for irinotecan.

Reference to nucleotides (or residues) may be according to their well known abbreviations. A "C" refers to a cytosine; "T" refers to "thymine"; "A" refers to adenine; and, "G" refers to guanine. If mRNA is used to determine a nucleotide sequence, "U" refers to uracil. In one study, the allele frequency for the variant allele (T) at position 3972 was 38.3% in Caucasians (n=100) and 27.3% African Americans (n=100). It is understood that a C is the most common nucleotide at position 3972. Because of that and the observations discussed herein, the activity of ABCC2 will be characterized relative to the activity of ABCC2 in persons with a C at 3972. Consequently, a normalized level of activity of ABCC2 in persons with a C at 3972 will be understood as a "normal level of ABCC2 activity." Moreover, in some embodiments of the invention, identification of a T at position 3972 on both alleles of the *ABCC2* gene is indicative of a lower than normal level of ABCC2 activity.

It will be understood that the term "determine" is used according to its ordinary and plain meaning to indicate "to ascertain definitely by observation, examination,

calculation, etc.,” according to the Oxford English Dictionary (2nd ed.). It will also be understood that the phrase “determining the sequence at position X” means that the nucleotide at that position is directly or indirectly identified. In some embodiments, the sequence at a particular position is determined, while in other embodiments, what is
5 determined at a particular position is that a particular nucleotide is *not* at that position.

Positions are indicated by conventional numbering where a negative sign (-) refers to nucleotides upstream (5') from the transcriptional start site (+1) (these sequences are in the promoter), unless otherwise designated. Sequences in the 5' untranslated region (5' UTR) may also be referred to using a negative sign, and in these cases, the positioning
10 is with respect to the translated portion, where the first nucleotide of a codon is understood as +1. Positions downstream of the translational start site may or may not have a plus sign (+). Furthermore, unless otherwise indicated or understood, identification of a position downstream of the transcriptional start site refers to a position with respect to only the coding region of the gene, that is, its exons and not the introns. In
15 some instances, positions within introns are referred to and the numbering for these positions is typically with respect to that intron alone, and not the gene as a whole.

It is contemplated that in methods of the invention, one or more sequences in one or both alleles of the *ABCC2* gene is determined. In some embodiments, both alleles of the patient are evaluated, while in others, only one allele is evaluated.

20 In further embodiments of the invention, methods also include obtaining a sample from a patient and using the sample to determine the sequence at position 3972. The sample may contain blood, serum, or a tissue biopsy, as well as buccal cells, mononuclear cells, or cancer cells.

Sequences may be determined by performing or conducting a hybridization assay,
25 an amplification assay, particularly one that is allele-specific, a sequencing or microsequencing assay.

The sequence at position 3972 may be determined directly or indirectly. A direct determination involves performing an assay with respect to that position. An indirect determination means that the sequence at position 3972 is determined based on data
30 regarding a different position, particularly by evaluating the sequence of a position in

linkage disequilibrium (LD) with a sequence at position 3972. In some embodiments, the sequence in LD with a sequence at position 3972 is in complete linkage disequilibrium with a sequence at 3972. In additional embodiments, the position in linkage disequilibrium with the sequence at position 3972 is selected from the group consisting of positions -1549 (promoter), -1019 (promoter), -24 (5' UTR), and +27 (intron 13). In some cases, more than one position in linkage disequilibrium with the sequence at position 3972 is evaluated to determine the sequence at position 3972. Therefore, in some embodiments of the invention, a haplotype that includes position 3972 is evaluated. In these embodiments, a determination of one or more sequences in one or both alleles of a gene in the haplotype is included in methods of the invention.

In methods of the invention, in some embodiments, an additional step of administering an *ABCC2* substrate to the patient is included. Likewise, in some embodiments, the step of administering an anticancer agent to the patient is included in methods of the invention. In some cases, the amount, formulation, or timing of the administration is based on the genotypic analysis of position 3972 of the *ABCC2* gene. In some embodiments of the invention, a patient is also provided additional anticancer therapy, such as the administration of a second anticancer agent or the performance of surgery on the patient. The second anticancer agent may be chemotherapy, particularly one that is not an *ABCC2* substrate or not the same *ABCC2* substrate that was already given to the patient, radiation therapy, immunotherapy, or gene therapy.

The present invention further concerns compositions that can be used to determine the sequence at position 3972 or any other sequence in LD with it. Accordingly, the present invention concerns kits for achieving methods of the invention. In some embodiments, the kits include one or more nucleic acids for determining the sequence at position 3972 in at one or both alleles of the *ABCC2* gene. In some embodiments, the nucleic acid is a primer for amplifying the sequence at position 3972 in the *ABCC2* gene. In others, the nucleic acid is a specific hybridization probe for detecting the sequence at position 3972 in the *ABCC2* gene. Additionally, it is contemplated that the specific hybridization probe can be comprised in an oligonucleotide array or microarray.

It is contemplated that any method or composition described herein can be implemented with respect to any other method or composition described herein. Similarly, any embodiment discussed with respect to one aspect of the invention may be used in the context of any other aspect of the invention.

5 Throughout this application, the term “about” is used to indicate that a value includes the standard deviation of error for the device or method being employed to determine the value.

10 The use of the word “a” or “an” when used in conjunction with the term “comprising” in the claims and/or the specification may mean “one,” but it is also consistent with the meaning of “one or more,” “at least one,” and “one or more than one.”

 The use of the term “or” in the claims is used to mean “and/or” unless explicitly indicated to refer to alternatives only or the alternative are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and “and/or.”

15 Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating specific embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

20

BRIEF DESCRIPTION OF THE DRAWINGS

25 The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

FIG. 1: ABCC2 3972C>T variant and AUC values of irinotecan and APC.

30 **FIG. 2:** ABCC2 3972C>T variant and AUC values of SN-38 and SN-38G.

DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The present invention provides improved methods and compositions for identifying the effects of polymorphisms in ABCC2 on the disposition of drugs and drug metabolites for the evaluation of the potential risk for drug toxicity or adverse events in an individual or patient. The development of these improved methods and compositions allows for the use of such an evaluation to optimize treatment of a patient and to lower the risk of toxicity or adverse events.

I. ABCC2

ABCC2, also referred to as MRP2 and cMOAT, functions as the major exporter of organic anions from the liver into the bile. In addition, ABCC2 is expressed on the apical membrane of epithelial cells such as enterocytes, renal proximal tubule epithelia, and gall bladder epithelia. ABCC2 is also expressed in some tumor tissues such as ovarian carcinoma, colorectal carcinoma, leukemia, mesothelioma, and hepatocarcinoma; and it has been suggested that tumor cells overexpressing ABCC2 acquire multidrug resistance (MDR) (Borst *et al.* (1999); Borst *et al.* (2000)).

ABCC2 is important from a pharmacological point of view because it is involved in the clearance of several clinically important drugs. One such drug is the anticancer drug irinotecan (CPT-11).

Irinotecan is also inactivated to oxidated metabolites (including APC) by CYP3A enzymes, and is activated to SN-38, which has a 100-1,000-fold higher antitumor activity than irinotecan, by carboxylesterase-2 (CES-2). SN-38 is glucuronidated by hepatic uridine diphosphate glucuronosyltransferases (UGTs) to form SN-38 glucuronide (10O-glucuronyl-SN-38, SN-38G), which is inactive and excreted into the bile and urine although, SN-38G might be deconjugated to form SN-38 by intestinal β -glucuronidase enzyme (Kaneda *et al.*, 1990). Irinotecan, SN-38, and SN-38G are known substrates for ABCC2. (Suzuki *et al.* (1999); Suzuki *et al.* (1998)).

The major dose-limiting toxicities of irinotecan include diarrhea and, to a lesser extent, myelosuppression. Irinotecan-induced diarrhea can be serious and often does not respond adequately to conventional antidiarrheal agents (Takasuna *et al.*, 1995). This diarrhea may be due to direct enteric injury caused by the active metabolite, SN-38,

which has been shown to accumulate in the intestine after intra peritoneal administration of irinotecan in athymic mice (Araki *et al.*, 1993).

It has been shown that there is an inverse relationship between SN-38 glucuronidation rates and severity of diarrheal incidences in patients treated with increasing doses of Irinotecan (Gupta *et al.*, 1994). These findings indicate that glucuronidation of SN-38 protects against Irinotecan-induced gastrointestinal toxicity. Therefore, differential rates of SN-38 glucuronidation among subjects may explain the considerable inter-individual variation in the pharmacokinetic parameter estimates and toxicities observed after treatment with anti-cancer drugs or exposure to xenobiotics (Gupta *et al.*, 1994; Gupta *et al.*, 1997).

The present invention demonstrates that the synonymous 3972C>T (exon 28) in ABCC2 is correlated with AUC (area under the curve) for irinotecan ($p=0.02$), APC ($p<0.0001$), APC/irinotecan ratio ($p<0.0001$), SN-38G ($p\leq 0.001$), and SN-38G/SN-38 ($p\leq 0.001$). Furthermore, the TT 3972 genotype was associated with higher AUC of irinotecan ($p=0.02$), APC ($p<0.0001$), and SN-38G ($p<0.0001$) compared to CT and CC patients. The phenotypic effect of 3972C>T was previously unknown, and identifies 3972C>T as a variant potentially affecting ABCC2 activity and suggests its biological function and clinical relevance for ABCC2 substrates. Thus, the present invention provides improved methods and compositions for evaluating the disposition of drugs and drug metabolites, and for evaluating the potential risk for drug toxicity in an individual or patient. The development of these improved methods and compositions allows for the use of such an evaluation to optimize treatment of a patient and to lower the risk of toxicity.

AUC is a measure of how much drug reaches the bloodstream in a set period of time. AUC is calculated by plotting drug blood concentration at various times over a specified period of time, usually 24 hours, and then measuring the area under the curve. AUC has a number of important uses in toxicology, biopharmaceutics, and pharmacokinetics. It is understood to be the time course or exposure of the patient to the drug.

The metabolism of irinotecan is merely illustrative of the present invention; the metabolism of other ABCC2 substrates is also contemplated. A summary of ABCC2 substrates is provided in Table 1 below. The table includes ABCC2 drug substrates.

5

Table 1. ABCC2 Substrates

Cysteinyl Leukotrienes LTC ₄ LTD ₄ LTE ₄ N-acetylated LTE ₄
GSH and GSH-Conjugates of Organic Compounds Reduced glutathione (GSH) Oxidized glutathione (GSSG) 2,4-dinitrophenol-S-glutathione Glutathione-bimane GSH Conjugate of bromosulphophthalein GSH Conjugate of bromoisovalerylurea GSH Conjugate of N-ethylmaleimide GSH Conjugate of ethacrynic acid GSH Conjugate of α -naphthylisothiocyanate GSH Conjugate of methylfluoroscein GSH Conjugate of prostaglandin A1 GSH Conjugate of (+)-anti-benzo[a]pyrene-7,8-diol-9,10-epoxide GSH Conjugate of 4-hydroxynonenal
GSH Conjugates of Metals Antimony Arsenic Bismuth Cadmium Copper Silver Zinc
Glucuronide Conjugates Bilirubin monoglucuronide Bilirubin diglucuronide 17 β estradiol 17 β -D-glucuronide Triiodothyronine-glucuronide p-nitrophenol- β -D-glucuronide 1-naphytol- β -D-glucoronide

E3040 glucuronide
 SN-38 glucuronide (SN-38G)
 Grepafloxacin glucuronide
 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanol glucuronide
 Telmisaltan glucuronide
 Acetaminophen glucuronide
 Diclofenac glucuronide
 Indomethacin glucuronide
 Glucuronide conjugates of 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine
 Liquiritigenin glucuronide
 Glycyrrhizin

Sulfated Conjugates

Dehydroepiandrosterone sulfate

Bile Salt Conjugates

Cholate-3-O-glucuronide
 Lithocholate-3-O-glucuronide
 Chenodeoxycholate-3-O-glucuronide
 Nordeoxycholate-3-O-glucuronide
 Nordeoxycholate-3-sulfate
 Lithocholate-3-sulfate
 Taurolithocholate-3-sulfate
 Glycolithocholate-3-sulfate
 Taurochenodeoxycholate-3-sulfate

Non-Conjugated Compounds

Bromosulphophthalein
 Dibromosulphophthalein
 Carboxyfluorescein
 Reduced folates
 Methotrexate
 CPT-11
 SN-38
 Ampicillin
 Ceftriaxone
 Cefodizime
 Grepafloxacin
 Pravastatin
 Temocaprilat
 BQ123
 p-aminohippuric acid
 Fluo-3
 Sulfinpyrazone (GSH coupled)

Vinblastine (GSH coupled) 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (GSH coupled) Etoposide Vincristine Doxorubicin Epirubicin Cisplatin

II. NUCLEIC ACIDS

Certain embodiments of the present invention concern various nucleic acids, including amplification primers, oligonucleotide probes, and other nucleic acid elements involved in the analysis of genomic DNA. In certain aspects, a nucleic acid comprises a wild-type, a mutant, or a polymorphic nucleic acid.

The term "nucleic acid" is well known in the art. A "nucleic acid" as used herein will generally refer to a molecule (*i.e.*, a strand) of DNA, RNA or a derivative or analog thereof, comprising a nucleobase. A nucleobase includes, for example, a naturally occurring purine or pyrimidine base found in DNA (*e.g.*, an adenine "A," a guanine "G," a thymine "T" or a cytosine "C") or RNA (*e.g.*, an A, a G, an uracil "U" or a C). The term "nucleic acid" encompass the terms "oligonucleotide" and "polynucleotide," each as a subgenus of the term "nucleic acid." The term "oligonucleotide" refers to a molecule of between about 3 and about 100 nucleobases in length. The term "polynucleotide" refers to at least one molecule of greater than about 100 nucleobases in length. A "gene" refers to coding sequence of a gene product, as well as introns and the promoter of the gene product. In addition to the *ABCC2* gene, other regulatory regions such as enhancers for *ABCC2* are contemplated as nucleic acids for use with compositions and methods of the claimed invention.

In some embodiments, nucleic acids of the invention comprise or are complementary to all or 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240,

250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 or more contiguous nucleotides of 1 (ABCC2 cDNA), SEQ ID NO:2 (ABCC2 exon 28), or SEQ ID NO:3.

These definitions generally refer to a single-stranded molecule, but in specific embodiments will also encompass an additional strand that is partially, substantially or fully complementary to the single-stranded molecule. Thus, a nucleic acid may encompass a double-stranded molecule or a triple-stranded molecule that comprises one or more complementary strand(s) or "complement(s)" of a particular sequence comprising a molecule. As used herein, a single stranded nucleic acid may be denoted by the prefix "ss", a double stranded nucleic acid by the prefix "ds", and a triple stranded nucleic acid by the prefix "ts."

In particular aspects, a nucleic acid encodes a protein, polypeptide, or peptide. In certain embodiments, the present invention concerns novel compositions comprising at least one proteinaceous molecule. As used herein, a "proteinaceous molecule," "proteinaceous composition," "proteinaceous compound," "proteinaceous chain," or "proteinaceous material" generally refers, but is not limited to, a protein of greater than about 200 amino acids or the full length endogenous sequence translated from a gene; a polypeptide of greater than about 100 amino acids; and/or a peptide of from about 3 to about 100 amino acids. All the "proteinaceous" terms described above may be used interchangeably herein.

1. Preparation of Nucleic Acids

A nucleic acid may be made by any technique known to one of ordinary skill in the art, such as for example, chemical synthesis, enzymatic production or biological production. Non-limiting examples of a synthetic nucleic acid (*e.g.*, a synthetic oligonucleotide), include a nucleic acid made by *in vitro* chemical synthesis using phosphotriester, phosphite or phosphoramidite chemistry and solid phase techniques such as described in European Patent 266,032, incorporated herein by reference, or via deoxynucleoside H-phosphonate intermediates as described by Froehler *et al.*, 1986 and

U.S. Patent 5,705,629, each incorporated herein by reference. In the methods of the present invention, one or more oligonucleotide may be used. Various different mechanisms of oligonucleotide synthesis have been disclosed in for example, U.S. Patents 4,659,774, 4,816,571, 5,141,813, 5,264,566, 4,959,463, 5,428,148, 5,554,744, 5,574,146, 5,602,244, each of which is incorporated herein by reference.

A non-limiting example of an enzymatically produced nucleic acid include one produced by enzymes in amplification reactions such as PCR™ (see for example, U.S. Patent 4,683,202 and U.S. Patent 4,682,195, each incorporated herein by reference), or the synthesis of an oligonucleotide described in U.S. Patent 5,645,897, incorporated herein by reference. A non-limiting example of a biologically produced nucleic acid includes a recombinant nucleic acid produced (*i.e.*, replicated) in a living cell, such as a recombinant DNA vector replicated in bacteria (see for example, Sambrook *et al.* 2001, incorporated herein by reference).

2. Purification of Nucleic Acids

A nucleic acid may be purified on polyacrylamide gels, cesium chloride centrifugation gradients, chromatography columns or by any other means known to one of ordinary skill in the art (see for example, Sambrook *et al.*, 2001, incorporated herein by reference). In some aspects, a nucleic acid is a pharmacologically acceptable nucleic acid. Pharmacologically acceptable compositions are known to those of skill in the art, and are described herein.

In certain aspects, the present invention concerns a nucleic acid that is an isolated nucleic acid. As used herein, the term "isolated nucleic acid" refers to a nucleic acid molecule (*e.g.*, an RNA or DNA molecule) that has been isolated free of, or is otherwise free of, the bulk of the total genomic and transcribed nucleic acids of one or more cells. In certain embodiments, "isolated nucleic acid" refers to a nucleic acid that has been isolated free of, or is otherwise free of, bulk of cellular components or *in vitro* reaction components such as for example, macromolecules such as lipids or proteins, small biological molecules, and the like.

3. Nucleic Acid Segments

In certain embodiments, the nucleic acid is a nucleic acid segment. As used herein, the term "nucleic acid segment," are fragments of a nucleic acid, such as, for a

non-limiting example, those that encode only part of a *ABCC2* gene locus or a *ABCC2* gene sequence. Thus, a "nucleic acid segment" may comprise any part of a gene sequence, including from about 2 nucleotides to the full length gene including promoter regions to the polyadenylation signal and any length that includes all the coding region.

5 Various nucleic acid segments may be designed based on a particular nucleic acid sequence, and may be of any length. By assigning numeric values to a sequence, for example, the first residue is 1, the second residue is 2, *etc.*, an algorithm defining all nucleic acid segments can be created:

$$n \text{ to } n + y$$

10 where *n* is an integer from 1 to the last number of the sequence and *y* is the length of the nucleic acid segment minus one, where *n + y* does not exceed the last number of the sequence. Thus, for a 10-mer, the nucleic acid segments correspond to bases 1 to 10, 2 to 11, 3 to 12 ... and so on. For a 15-mer, the nucleic acid segments correspond to bases 1 to 15, 2 to 16, 3 to 17 ... and so on. For a 20-mer, the nucleic segments correspond to bases
15 1 to 20, 2 to 21, 3 to 22 ... and so on. In certain embodiments, the nucleic acid segment may be a probe or primer. As used herein, a "probe" generally refers to a nucleic acid used in a detection method or composition. As used herein, a "primer" generally refers to a nucleic acid used in an extension or amplification method or composition.

4. Nucleic Acid Complements

20 The present invention also encompasses a nucleic acid that is complementary to a nucleic acid. A nucleic acid is "complement(s)" or is "complementary" to another nucleic acid when it is capable of base-pairing with another nucleic acid according to the standard Watson-Crick, Hoogsteen or reverse Hoogsteen binding complementarity rules. As used herein "another nucleic acid" may refer to a separate molecule or a spatial
25 separated sequence of the same molecule. In preferred embodiments, a complement is a hybridization probe or amplification primer for the detection of a nucleic acid polymorphism.

 As used herein, the term "complementary" or "complement" also refers to a nucleic acid comprising a sequence of consecutive nucleobases or semiconsecutive
30 nucleobases (*e.g.*, one or more nucleobase moieties are not present in the molecule) capable of hybridizing to another nucleic acid strand or duplex even if less than all the

nucleobases do not base pair with a counterpart nucleobase. However, in some diagnostic or detection embodiments, completely complementary nucleic acids are preferred.

5 **III. NUCLEIC ACID DETECTION**

Some embodiments of the invention concern identifying polymorphisms in *ABCC2*, correlating genotype or haplotype to phenotype, wherein the phenotype is altered *ABCC2* activity or expression, and then identifying such polymorphisms in patients who have or will be given irinotecan or other drugs or compounds that are

10 *ABCC2* substrates. Thus, the present invention involves assays for identifying polymorphisms and other nucleic acid detection methods. Nucleic acids, therefore, have utility as probes or primers for embodiments involving nucleic acid hybridization. They may be used in diagnostic or screening methods of the present invention. Detection of nucleic acids encoding *ABCC2*, as well as nucleic acids involved in the expression or

15 stability of *ABCC2* polypeptides or transcripts, are encompassed by the invention. General methods of nucleic acid detection methods are provided below, followed by specific examples employed for the identification of polymorphisms, including single nucleotide polymorphisms (SNPs).

A. Hybridization

20 The use of a probe or primer of between 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 50, 60, 70, 80, 90, or 100 nucleotides, preferably between 17 and 100 nucleotides in length, or in some aspects of the invention up to 1-2 kilobases or more in length, allows the formation of a duplex molecule that is both stable and selective. Molecules having complementary sequences over contiguous stretches greater

25 than 20 bases in length are generally preferred, to increase stability and/or selectivity of the hybrid molecules obtained. One will generally prefer to design nucleic acid molecules for hybridization having one or more complementary sequences of 20 to 30 nucleotides, or even longer where desired. Such fragments may be readily prepared, for example, by directly synthesizing the fragment by chemical means or by introducing

30 selected sequences into recombinant vectors for recombinant production.

Accordingly, the nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of DNAs and/or RNAs or to provide primers for amplification of DNA or RNA from samples. Depending on the application envisioned, one would desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of the probe or primers for the target sequence.

For applications requiring high selectivity, one will typically desire to employ relatively high stringency conditions to form the hybrids. For example, relatively low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.10 M NaCl at temperatures of about 50°C to about 70°C. Such high stringency conditions tolerate little, if any, mismatch between the probe or primers and the template or target strand and would be particularly suitable for isolating specific genes or for detecting a specific polymorphism. It is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide. For example, under highly stringent conditions, hybridization to filter-bound DNA may be carried out in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1 x SSC/0.1% SDS at 68°C (Ausubel *et al.*, 1989).

Conditions may be rendered less stringent by increasing salt concentration and/or decreasing temperature. For example, a medium stringency condition could be provided by about 0.1 to 0.25M NaCl at temperatures of about 37°C to about 55°C, while a low stringency condition could be provided by about 0.15M to about 0.9M salt, at temperatures ranging from about 20°C to about 55°C. Under low stringent conditions, such as moderately stringent conditions the washing may be carried out for example in 0.2 x SSC/0.1% SDS at 42°C (Ausubel *et al.*, 1989). Hybridization conditions can be readily manipulated depending on the desired results.

In other embodiments, hybridization may be achieved under conditions of, for example, 50mM Tris-HCl (pH 8.3), 75mM KCl, 3mM MgCl₂, 1.0mM dithiothreitol, at temperatures between approximately 20°C to about 37°C. Other hybridization conditions utilized could include approximately 10mM Tris-HCl (pH 8.3), 50mM KCl, 1.5mM MgCl₂, at temperatures ranging from approximately 40°C to about 72°C.

In certain embodiments, it will be advantageous to employ nucleic acids of defined sequences of the present invention in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of being detected. In preferred
5 embodiments, one may desire to employ a fluorescent label or an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmentally undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known that can be employed to provide a detection means that is visibly or
10 spectrophotometrically detectable, to identify specific hybridization with complementary nucleic acid containing samples. In other aspects, a particular nuclease cleavage site may be present and detection of a particular nucleotide sequence can be determined by the presence or absence of nucleic acid cleavage.

In general, it is envisioned that the probes or primers described herein will be
15 useful as reagents in solution hybridization, as in PCR, for detection of expression or genotype of corresponding genes, as well as in embodiments employing a solid phase. In embodiments involving a solid phase, the test DNA (or RNA) is adsorbed or otherwise affixed to a selected matrix or surface. This fixed, single-stranded nucleic acid is then subjected to hybridization with selected probes under desired conditions. The conditions
20 selected will depend on the particular circumstances (depending, for example, on the G+C content, type of target nucleic acid, source of nucleic acid, size of hybridization probe, *etc.*). Optimization of hybridization conditions for the particular application of interest is well known to those of skill in the art. After washing of the hybridized molecules to remove non-specifically bound probe molecules, hybridization is detected,
25 and/or quantified, by determining the amount of bound label. Representative solid phase hybridization methods are disclosed in U.S. Patents 5,843,663, 5,900,481 and 5,919,626. Other methods of hybridization that may be used in the practice of the present invention are disclosed in U.S. Patents 5,849,481, 5,849,486 and 5,851,772. The relevant portions of these and other references identified in this section of the Specification are
30 incorporated herein by reference.

B. Amplification of Nucleic Acids

Nucleic acids used as a template for amplification may be isolated from cells, tissues or other samples according to standard methodologies (Sambrook *et al.*, 2001). In certain embodiments, analysis is performed on whole cell or tissue homogenates or biological fluid samples with or without substantial purification of the template nucleic acid. The nucleic acid may be genomic DNA or fractionated or whole cell RNA. Where RNA is used, it may be desired to first convert the RNA to a complementary DNA.

The term "primer," as used herein, is meant to encompass any nucleic acid that is capable of priming the synthesis of a nascent nucleic acid in a template-dependent process. Typically, primers are oligonucleotides from ten to twenty and/or thirty base pairs in length, but longer sequences can be employed. Primers may be provided in double-stranded and/or single-stranded form, although the single-stranded form is preferred.

Pairs of primers designed to selectively hybridize to nucleic acids corresponding to the *ABCC2* gene locus (GenBank accession NT030059, incorporated herein by reference), or variants thereof, and fragments thereof are contacted with the template nucleic acid under conditions that permit selective hybridization. Depending upon the desired application, high stringency hybridization conditions may be selected that will only allow hybridization to sequences that are completely complementary to the primers. In other embodiments, hybridization may occur under reduced stringency to allow for amplification of nucleic acids that contain one or more mismatches with the primer sequences. Once hybridized, the template-primer complex is contacted with one or more enzymes that facilitate template-dependent nucleic acid synthesis. Multiple rounds of amplification, also referred to as "cycles," are conducted until a sufficient amount of amplification product is produced.

The amplification product may be detected, analyzed or quantified. In certain applications, the detection may be performed by visual means. In certain applications, the detection may involve indirect identification of the product via chemiluminescence, radioactive scintigraphy of incorporated radiolabel or fluorescent label or even via a system using electrical and/or thermal impulse signals (Affymax technology; Bellus, 1994).

A number of template dependent processes are available to amplify the oligonucleotide sequences present in a given template sample. One of the best known amplification methods is the polymerase chain reaction (referred to as PCR™) which is described in detail in U.S. Patents 4,683,195, 4,683,202 and 4,800,159, and in Innis *et al.*,
5 1988, each of which is incorporated herein by reference in their entirety.

Another method for amplification is ligase chain reaction ("LCR"), disclosed in European Application No. 320 308, incorporated herein by reference in its entirety. U.S. Patent 4,883,750 describes a method similar to LCR for binding probe pairs to a target sequence. A method based on PCR™ and oligonucleotide ligase assay (OLA) (described
10 in further detail below), disclosed in U.S. Patent 5,912,148, may also be used.

Alternative methods for amplification of target nucleic acid sequences that may be used in the practice of the present invention are disclosed in U.S. Patents 5,843,650, 5,846,709, 5,846,783, 5,849,546, 5,849,497, 5,849,547, 5,858,652, 5,866,366, 5,916,776, 5,922,574, 5,928,905, 5,928,906, 5,932,451, 5,935,825, 5,939,291 and 5,942,391, Great
15 Britain Application 2 202 328, and in PCT Application PCT/US89/01025, each of which is incorporated herein by reference in its entirety. Qbeta Replicase, described in PCT Application PCT/US87/00880, may also be used as an amplification method in the present invention.

An isothermal amplification method, in which restriction endonucleases and
20 ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[alpha-thio]-triphosphates in one strand of a restriction site may also be useful in the amplification of nucleic acids in the present invention (Walker *et al.*, 1992). Strand Displacement Amplification (SDA), disclosed in U.S. Patent 5,916,779, is another method of carrying out isothermal amplification of nucleic acids which involves multiple
25 rounds of strand displacement and synthesis, *i.e.*, nick translation

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS), including nucleic acid sequence based amplification (NASBA) and 3SR (Kwoh *et al.*, 1989; PCT Application WO 88/10315, incorporated
herein by reference in their entirety). European Application 329 822 disclose a nucleic
30 acid amplification process involving cyclically synthesizing single-stranded RNA

("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention.

PCT Application WO 89/06700 (incorporated herein by reference in its entirety) disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter region/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic, *i.e.*, new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" and "one-sided PCR" (Frohman, 1990; Ohara *et al.*, 1989).

C. Detection of Nucleic Acids

Following any amplification, it may be desirable to separate the amplification product from the template and/or the excess primer. In one embodiment, amplification products are separated by agarose, agarose-acrylamide or polyacrylamide gel electrophoresis using standard methods (Sambrook *et al.*, 2001). Separated amplification products may be cut out and eluted from the gel for further manipulation. Using low melting point agarose gels, the separated band may be removed by heating the gel, followed by extraction of the nucleic acid.

Separation of nucleic acids may also be effected by spin columns and/or chromatographic techniques known in art. There are many kinds of chromatography which may be used in the practice of the present invention, including adsorption, partition, ion-exchange, hydroxylapatite, molecular sieve, reverse-phase, column, paper, thin-layer, and gas chromatography as well as HPLC.

In certain embodiments, the amplification products are visualized, with or without separation. A typical visualization method involves staining of a gel with ethidium bromide and visualization of bands under UV light. Alternatively, if the amplification products are integrally labeled with radio- or fluorometrically-labeled nucleotides, the separated amplification products can be exposed to x-ray film or visualized under the appropriate excitatory spectra.

In one embodiment, following separation of amplification products, a labeled nucleic acid probe is brought into contact with the amplified marker sequence. The probe preferably is conjugated to a chromophore but may be radiolabeled. In another

embodiment, the probe is conjugated to a binding partner, such as an antibody or biotin, or another binding partner carrying a detectable moiety.

In particular embodiments, detection is by Southern blotting and hybridization with a labeled probe. The techniques involved in Southern blotting are well known to those of skill in the art (see Sambrook *et al.*, 2001). One example of the foregoing is described in U.S. Patent 5,279,721, incorporated by reference herein, which discloses an apparatus and method for the automated electrophoresis and transfer of nucleic acids. The apparatus permits electrophoresis and blotting without external manipulation of the gel and is ideally suited to carrying out methods according to the present invention.

Other methods of nucleic acid detection that may be used in the practice of the instant invention are disclosed in U.S. Patents 5,840,873, 5,843,640, 5,843,651, 5,846,708, 5,846,717, 5,846,726, 5,846,729, 5,849,487, 5,853,990, 5,853,992, 5,853,993, 5,856,092, 5,861,244, 5,863,732, 5,863,753, 5,866,331, 5,905,024, 5,910,407, 5,912,124, 5,912,145, 5,919,630, 5,925,517, 5,928,862, 5,928,869, 5,929,227, 5,932,413 and 5,935,791, each of which is incorporated herein by reference.

D. Other Assays

Other methods for genetic screening may be used within the scope of the present invention, for example, to detect mutations in genomic DNA, cDNA and/or RNA samples. Methods used to detect point mutations include denaturing gradient gel electrophoresis ("DGGE"), restriction fragment length polymorphism analysis ("RFLP"), chemical or enzymatic cleavage methods, direct sequencing of target regions amplified by PCR™ (see above), single-strand conformation polymorphism analysis ("SSCP") and other methods well known in the art.

One method of screening for point mutations is based on RNase cleavage of base pair mismatches in RNA/DNA or RNA/RNA heteroduplexes. As used herein, the term "mismatch" is defined as a region of one or more unpaired or mispaired nucleotides in a double-stranded RNA/RNA, RNA/DNA or DNA/DNA molecule. This definition thus includes mismatches due to insertion/deletion mutations, as well as single or multiple base point mutations.

U.S. Patent 4,946,773 describes an RNase A mismatch cleavage assay that involves annealing single-stranded DNA or RNA test samples to an RNA probe, and

subsequent treatment of the nucleic acid duplexes with RNase A. For the detection of mismatches, the single-stranded products of the RNase A treatment, electrophoretically separated according to size, are compared to similarly treated control duplexes. Samples containing smaller fragments (cleavage products) not seen in the control duplex are scored as positive.

Other investigators have described the use of RNase I in mismatch assays. The use of RNase I for mismatch detection is described in literature from Promega Biotech. Promega markets a kit containing RNase I that is reported to cleave three out of four known mismatches. Others have described using the MutS protein or other DNA-repair enzymes for detection of single-base mismatches.

Alternative methods for detection of deletion, insertion or substitution mutations that may be used in the practice of the present invention are disclosed in U.S. Patents 5,849,483, 5,851,770, 5,866,337, 5,925,525 and 5,928,870, each of which is incorporated herein by reference in its entirety.

E. Specific Examples of SNP Screening Methods

Spontaneous mutations that arise during the course of evolution in the genomes of organisms are often not immediately transmitted throughout all of the members of the species, thereby creating polymorphic alleles that co-exist in the species populations. Often polymorphisms are the cause of genetic diseases. Several classes of polymorphisms have been identified. For example, variable nucleotide type polymorphisms (VNTRs), arise from spontaneous tandem duplications of di- or trinucleotide repeated motifs of nucleotides. If such variations alter the lengths of DNA fragments generated by restriction endonuclease cleavage, the variations are referred to as restriction fragment length polymorphisms (RFLPs). RFLPs are been widely used in human and animal genetic analyses.

Another class of polymorphisms are generated by the replacement of a single nucleotide. Such single nucleotide polymorphisms (SNPs) rarely result in changes in a restriction endonuclease site. Thus, SNPs are rarely detectable restriction fragment length analysis. SNPs are the most common genetic variations and occur once every 100 to 300 bases and several SNP mutations have been found that affect a single nucleotide in a protein-encoding gene in a manner sufficient to actually cause a genetic disease. SNP

diseases are exemplified by hemophilia, sickle-cell anemia, hereditary hemochromatosis, late-onset alzheimer disease *etc.*

In context of the present invention, polymorphic mutations that affect the activity and/or level of the *ABCC2* gene product, which is responsible for the transport of numerous compounds across cell membranes, will be determined by a series of screening methods. To do this, a sample (such as blood or other bodily fluid or tissue sample) will be taken from a patient for genotype analysis. The presence or absence of SNPs will determine the ability of the screened individuals to metabolize irinotecan and other agents that are transported by *ABCC2*. According to methods provided by the invention, these results will be used to adjust and/or alter the dose of irinotecan or other agent administered to an individual in order to reduce drug side effects. In one embodiment, the presence of the 3972C>T variant in the *ABCC2* gene will be determined. The identification of a T at position 3972 on both alleles would indicate that the patient will be slower to dispose of *ABCC2* substrates (*e.g.*, irinotecan) than a patient with a C at position 3972 on one or both alleles. Thus, to minimize drug toxicity, it may be desirable to administer a lower drug dose to the patient having a T at position 3972 on both alleles.

In some embodiments, the methods and compositions of the present invention involve determining the sequence at polymorphic sites in linkage disequilibrium with the sequence at position 3972 of the *ABCC2* gene. For example, a common haplotype with the 3972 variant is one that includes two promoter variants (-1549A>G and -1019A>G) and a 5' UTR variant (-24C>T). Another haplotype including the 3972 variant and the -1549 and -1019 promoter variants is also common. Thus, in certain embodiments, the methods and compositions of the present invention comprise detecting one or more of the -1549A>G, -1019A>G, or -24C>T variants in the *ABCC2* gene. Yet another haplotype with the 3972 variant includes the -1549A>G promoter variant and an intronic variant in intron 13 (+27C>G). Thus, in certain embodiments, the methods and compositions of the present invention comprise detecting one or both of the -1549A>G or +27C>G variants in the *ABCC2* gene.

SNPs can be the result of deletions, point mutations and insertions and in general any single base alteration, whatever the cause, can result in a SNP. The greater frequency of SNPs means that they can be more readily identified than the other classes of

polymorphisms. The greater uniformity of their distribution permits the identification of SNPs "nearer" to a particular trait of interest. The combined effect of these two attributes makes SNPs extremely valuable. For example, if a particular trait (*e.g.*, inability to efficiently metabolize irinotecan) reflects a mutation at a particular locus, then any polymorphism that is linked to the particular locus can be used to predict the probability that an individual will exhibit that trait.

Several methods have been developed to screen polymorphisms and some examples are listed below. The reference of Kwok and Chen (2003) and Kwok (2001) provide overviews of some of these methods; both of these references are specifically incorporated by reference.

SNPs relating to ABCC2 can be characterized by the use of any of these methods or suitable modification thereof. Such methods include the direct or indirect sequencing of the site, the use of restriction enzymes where the respective alleles of the site create or destroy a restriction site, the use of allele-specific hybridization probes, the use of antibodies that are specific for the proteins encoded by the different alleles of the polymorphism, or any other biochemical interpretation.

i) DNA Sequencing

The most commonly used method of characterizing a polymorphism is direct DNA sequencing of the genetic locus that flanks and includes the polymorphism. Such analysis can be accomplished using either the "dideoxy-mediated chain termination method," also known as the "Sanger Method" (Sanger *et al.*, 1975) or the "chemical degradation method," also known as the "Maxam-Gilbert method" (Maxam *et al.*, 1977). Sequencing in combination with genomic sequence-specific amplification technologies, such as the polymerase chain reaction may be utilized to facilitate the recovery of the desired genes (Mullis *et al.*, 1986; European Patent Application 50,424; European Patent Application. 84,796, European Patent Application 258,017, European Patent Application. 237,362; European Patent Application. 201,184; U.S. Patents 4,683,202; 4,582,788; and 4,683,194), all of the above incorporated herein by reference.

ii) Exonuclease Resistance

Other methods that can be employed to determine the identity of a nucleotide present at a polymorphic site utilize a specialized exonuclease-resistant nucleotide

derivative (U.S. Patent. 4,656,127). A primer complementary to an allelic sequence immediately 3'-to the polymorphic site is hybridized to the DNA under investigation. If the polymorphic site on the DNA contains a nucleotide that is complementary to the particular exonuclease-resistant nucleotide derivative present, then that derivative will be incorporated by a polymerase onto the end of the hybridized primer. Such incorporation makes the primer resistant to exonuclease cleavage and thereby permits its detection. As the identity of the exonuclease-resistant derivative is known one can determine the specific nucleotide present in the polymorphic site of the DNA.

iii) Microsequencing Methods

Several other primer-guided nucleotide incorporation procedures for assaying polymorphic sites in DNA have been described (Komher *et al.*, 1989; Sokolov, 1990; Syvanen 1990; Kuppuswamy *et al.*, 1991; Prezant *et al.*, 1992; Ugozzoli *et al.*, 1992; Nyren *et al.*, 1993). These methods rely on the incorporation of labeled deoxynucleotides to discriminate between bases at a polymorphic site. As the signal is proportional to the number of deoxynucleotides incorporated, polymorphisms that occur in runs of the same nucleotide result in a signal that is proportional to the length of the run (Syvanen *et al.*, 1990).

iv) Extension in Solution

French Patent 2,650,840 and PCT Application WO91/02087 discuss a solution-based method for determining the identity of the nucleotide of a polymorphic site. According to these methods, a primer complementary to allelic sequences immediately 3'-to a polymorphic site is used. The identity of the nucleotide of that site is determined using labeled dideoxynucleotide derivatives which are incorporated at the end of the primer if complementary to the nucleotide of the polymorphic site.

v) Genetic Bit Analysis or Solid-Phase Extension

PCT Application WO92/15712 describes a method that uses mixtures of labeled terminators and a primer that is complementary to the sequence 3' to a polymorphic site. The labeled terminator that is incorporated is complementary to the nucleotide present in the polymorphic site of the target molecule being evaluated and is thus identified. Here the primer or the target molecule is immobilized to a solid phase.

vi) Oligonucleotide Ligation Assay (OLA)

This is another solid phase method that uses different methodology (Landegren *et al.*, 1988). Two oligonucleotides, capable of hybridizing to abutting sequences of a single strand of a target DNA are used. One of these oligonucleotides is biotinylated while the other is detectably labeled. If the precise complementary sequence is found in a target molecule, the oligonucleotides will hybridize such that their termini abut, and create a ligation substrate. Ligation permits the recovery of the labeled oligonucleotide by using avidin. Other nucleic acid detection assays, based on this method, combined with PCR have also been described (Nickerson *et al.*, 1990). Here PCR is used to achieve the exponential amplification of target DNA, which is then detected using the OLA.

vii) Ligase/Polymerase-Mediated Genetic Bit Analysis

U.S. Patent 5,952,174 describes a method that also involves two primers capable of hybridizing to abutting sequences of a target molecule. The hybridized product is formed on a solid support to which the target is immobilized. Here the hybridization occurs such that the primers are separated from one another by a space of a single nucleotide. Incubating this hybridized product in the presence of a polymerase, a ligase, and a nucleoside triphosphate mixture containing at least one deoxynucleoside triphosphate allows the ligation of any pair of abutting hybridized oligonucleotides. Addition of a ligase results in two events required to generate a signal, extension and ligation. This provides a higher specificity and lower "noise" than methods using either extension or ligation alone and unlike the polymerase-based assays, this method enhances the specificity of the polymerase step by combining it with a second hybridization and a ligation step for a signal to be attached to the solid phase.

viii) Other Methods To Detect SNPs

Several other specific methods for SNP detection and identification are presented below and may be used as such or with suitable modifications in conjunction with identifying polymorphisms of the *ABCC2* gene in the present invention. Several other methods are also described on the SNP web site of the NCBI at the website www.ncbi.nlm.nih.gov/SNP, incorporated herein by reference.

In a particular embodiment, extended haplotypes may be determined at any given locus in a population, which allows one to identify exactly which SNPs will be redundant and which will be essential in association studies. The latter is referred to as 'haplotype tag SNPs (htSNPs)', markers that capture the haplotypes of a gene or a region of linkage disequilibrium. See Johnson *et al.* (2001) and Ke and Cardon (2003), each of which is
5 incorporated herein by reference, for exemplary methods.

The VDA-assay utilizes PCR amplification of genomic segments by long PCR methods using TaKaRa LA Taq reagents and other standard reaction conditions. The long amplification can amplify DNA sizes of about 2,000-12,000 bp. Hybridization of
10 products to variant detector array (VDA) can be performed by a Affymetrix High Throughput Screening Center and analyzed with computerized software.

A method called Chip Assay uses PCR amplification of genomic segments by standard or long PCR protocols. Hybridization products are analyzed by VDA, Halushka *et al.* (1999), incorporated herein by reference. SNPs are generally classified as "Certain" or "Likely" based on computer analysis of hybridization patterns. By comparison to
15 alternative detection methods such as nucleotide sequencing, "Certain" SNPs have been confirmed 100% of the time; and "Likely" SNPs have been confirmed 73% of the time by this method.

Other methods simply involve PCR amplification following digestion with the relevant restriction enzyme. Yet others involve sequencing of purified PCR products
20 from known genomic regions.

In yet another method, individual exons or overlapping fragments of large exons are PCR-amplified. Primers are designed from published or database sequences and PCR-amplification of genomic DNA is performed using the following conditions: 200 ng
25 DNA template, 0.5 μ M each primer, 80 μ M each of dCTP, dATP, dTTP and dGTP, 5% formamide, 1.5mM MgCl₂, 0.5U of Taq polymerase and 0.1 volume of the Taq buffer. Thermal cycling is performed and resulting PCR-products are analyzed by PCR-single strand conformation polymorphism (PCR-SSCP) analysis, under a variety of conditions, *e.g.* 5 or 10% polyacrylamide gel with 15% urea, with or without 5% glycerol.
30 Electrophoresis is performed overnight. PCR-products that show mobility shifts are reamplified and sequenced to identify nucleotide variation.

In a method called CGAP-GAI (DEMIGLACE), sequence and alignment data (from a PHRAP.ace file), quality scores for the sequence base calls (from PHRED quality files), distance information (from PHYLIP dnadist and neighbour programs) and base-calling data (from PHRED '-d' switch) are loaded into memory. Sequences are aligned and examined for each vertical chunk ('slice') of the resulting assembly for disagreement. Any such slice is considered a candidate SNP (DEMIGLACE). A number of filters are used by DEMIGLACE to eliminate slices that are not likely to represent true polymorphisms. These include filters that: (i) exclude sequences in any given slice from SNP consideration where neighboring sequence quality scores drop 40% or more; (ii) exclude calls in which peak amplitude is below the fifteenth percentile of all base calls for that nucleotide type; (iii) disqualify regions of a sequence having a high number of disagreements with the consensus from participating in SNP calculations; (iv) removed from consideration any base call with an alternative call in which the peak takes up 25% or more of the area of the called peak; (v) exclude variations that occur in only one read direction. PHRED quality scores were converted into probability-of-error values for each nucleotide in the slice. Standard Bayesian methods are used to calculate the posterior probability that there is evidence of nucleotide heterogeneity at a given location.

In a method called CU-RDF (RESEQ), PCR amplification is performed from DNA isolated from blood using specific primers for each SNP, and after typical cleanup protocols to remove unused primers and free nucleotides, direct sequencing using the same or nested primers.

In a method called DEBNICK (METHOD-B), a comparative analysis of clustered EST sequences is performed and confirmed by fluorescent-based DNA sequencing. In a related method, called DEBNICK (METHOD-C), comparative analysis of clustered EST sequences with phred quality > 20 at the site of the mismatch, average phred quality >= 20 over 5 bases 5'-FLANK and 3' to the SNP, no mismatches in 5 bases 5' and 3' to the SNP, at least two occurrences of each allele is performed and confirmed by examining traces.

In a method identified by ERO (RESEQ), new primers sets are designed for electronically published STSs and used to amplify DNA from 10 different mouse strains. The amplification product from each strain is then gel purified and sequenced using a

standard dideoxy, cycle sequencing technique with ^{33}P -labeled terminators. All the ddATP terminated reactions are then loaded in adjacent lanes of a sequencing gel followed by all of the ddGTP reactions and so on. SNPs are identified by visually scanning the radiographs.

5 In another method identified as ERO (RESEQ-HT), new primers sets are designed for electronically published murine DNA sequences and used to amplify DNA from 10 different mouse strains. The amplification product from each strain is prepared for sequencing by treating with Exonuclease I and Shrimp Alkaline Phosphatase. Sequencing is performed using ABI Prism Big Dye Terminator Ready Reaction Kit
10 (Perkin-Elmer) and sequence samples are run on the 3700 DNA Analyzer (96 Capillary Sequencer).

 FGU-CBT (SCA2-SNP) identifies a method where the region containing the SNP were PCR amplified using the primers SCA2-FP3 and SCA2-RP3. Approximately 100 ng of genomic DNA is amplified in a 50 ml reaction volume containing a final
15 concentration of 5mM Tris, 25mM KCl, 0.75mM MgCl_2 , 0.05% gelatin, 20pmol of each primer and 0.5U of Taq DNA polymerase. Samples are denatured, annealed and extended and the PCR product is purified from a band cut out of the agarose gel using, for example, the QIAquick gel extraction kit (Qiagen) and is sequenced using dye terminator chemistry on an ABI Prism 377 automated DNA sequencer with the PCR
20 primers.

 In a method identified as JBLACK (SEQ/RESTRICT), two independent PCR reactions are performed with genomic DNA. Products from the first reaction are analyzed by sequencing, indicating a unique FspI restriction site. The mutation is confirmed in the product of the second PCR reaction by digesting with Fsp I.

25 In a method described as KWOK(1), SNPs are identified by comparing high quality genomic sequence data from four randomly chosen individuals by direct DNA sequencing of PCR products with dye-terminator chemistry (see Kwok *et al.*, 1996). In a related method identified as KWOK(2) SNPs are identified by comparing high quality genomic sequence data from overlapping large-insert clones such as bacterial artificial
30 chromosomes (BACs) or P1-based artificial chromosomes (PACs). An STS containing this SNP is then developed and the existence of the SNP in various populations is

confirmed by pooled DNA sequencing (see Taillon-Miller *et al.*, 1998). In another similar method called KWOK(3), SNPs are identified by comparing high quality genomic sequence data from overlapping large-insert clones BACs or PACs. The SNPs found by this approach represent DNA sequence variations between the two donor chromosomes but the allele frequencies in the general population have not yet been determined. In method KWOK(5), SNPs are identified by comparing high quality genomic sequence data from a homozygous DNA sample and one or more pooled DNA samples by direct DNA sequencing of PCR products with dye-terminator chemistry. The STSs used are developed from sequence data found in publicly available databases. Specifically, these STSs are amplified by PCR against a complete hydatidiform mole (CHM) that has been shown to be homozygous at all loci and a pool of DNA samples from 80 CEPH parents (see Kwok *et al.*, 1994).

In another such method, KWOK (OverlapSnmpDetectionWithPolyBayes), SNPs are discovered by automated computer analysis of overlapping regions of large-insert human genomic clone sequences. For data acquisition, clone sequences are obtained directly from large-scale sequencing centers. This is necessary because base quality sequences are not present/available through GenBank. Raw data processing involves analyzed of clone sequences and accompanying base quality information for consistency. Finished ('base perfect', error rate lower than 1 in 10,000 bp) sequences with no associated base quality sequences are assigned a uniform base quality value of 40 (1 in 10,000 bp error rate). Draft sequences without base quality values are rejected. Processed sequences are entered into a local database. A version of each sequence with known human repeats masked is also stored. Repeat masking is performed with the program "MASKERAID." Overlap detection: Putative overlaps are detected with the program "WUBLAST." Several filtering steps followed in order to eliminate false overlap detection results, *i.e.* similarities between a pair of clone sequences that arise due to sequence duplication as opposed to true overlap. Total length of overlap, overall percent similarity, number of sequence differences between nucleotides with high base quality value "high-quality mismatches." Results are also compared to results of restriction fragment mapping of genomic clones at Washington University Genome Sequencing Center, finisher's reports on overlaps, and results of the sequence contig

building effort at the NCBI. SNP detection: Overlapping pairs of clone sequence are analyzed for candidate SNP sites with the 'POLYBAYES' SNP detection software. Sequence differences between the pair of sequences are scored for the probability of representing true sequence variation as opposed to sequencing error. This process requires the presence of base quality values for both sequences. High-scoring candidates are extracted. The search is restricted to substitution-type single base pair variations. Confidence score of candidate SNP is computed by the POLYBAYES software.

In method identified by KWOK (TaqMan assay), the TaqMan assay is used to determine genotypes for 90 random individuals. In method identified by KYUGEN(Q1), DNA samples of indicated populations are pooled and analyzed by PLACE-SSCP. Peak heights of each allele in the pooled analysis are corrected by those in a heterozygote, and are subsequently used for calculation of allele frequencies. Allele frequencies higher than 10% are reliably quantified by this method. Allele frequency = 0 (zero) means that the allele was found among individuals, but the corresponding peak is not seen in the examination of pool. Allele frequency = 0-0.1 indicates that minor alleles are detected in the pool but the peaks are too low to reliably quantify.

In yet another method identified as KYUGEN (Method1), PCR products are post-labeled with fluorescent dyes and analyzed by an automated capillary electrophoresis system under SSCP conditions (PLACE-SSCP). Four or more individual DNAs are analyzed with or without two pooled DNA (Japanese pool and CEPH parents pool) in a series of experiments. Alleles are identified by visual inspection. Individual DNAs with different genotypes are sequenced and SNPs identified. Allele frequencies are estimated from peak heights in the pooled samples after correction of signal bias using peak heights in heterozygotes. For the PCR primers are tagged to have 5'-ATT or 5'-GTT at their ends for post-labeling of both strands. Samples of DNA (10 ng/ul) are amplified in reaction mixtures containing the buffer (10mM Tris-HCl, pH 8.3 or 9.3, 50mM KCl, 2.0mM MgCl₂), 0.25μM of each primer, 200μM of each dNTP, and 0.025 units/μl of Taq DNA polymerase premixed with anti-Taq antibody. The two strands of PCR products are differentially labeled with nucleotides modified with R110 and R6G by an exchange reaction of Klenow fragment of DNA polymerase I. The reaction is stopped by adding EDTA, and unincorporated nucleotides are dephosphorylated by adding calf intestinal

alkaline phosphatase. For the SSCP: an aliquot of fluorescently labeled PCR products and TAMRA-labeled internal markers are added to deionized formamide, and denatured. Electrophoresis is performed in a capillary using an ABI Prism 310 Genetic Analyzer. Genescan softwares (P-E Biosystems) are used for data collection and data processing.

5 DNA of individuals (two to eleven) including those who showed different genotypes on SSCP are subjected for direct sequencing using big-dye terminator chemistry, on ABI Prism 310 sequencers. Multiple sequence trace files obtained from ABI Prism 310 are processed and aligned by Phred/Phrap and viewed using Consed viewer. SNPs are identified by PolyPhred software and visual inspection.

10 In yet another method identified as KYUGEN (Method2), individuals with different genotypes are searched by denaturing HPLC (DHPLC) or PLACE-SSCP (Inazuka *et al.*, 1997) and their sequences are determined to identify SNPs. PCR is performed with primers tagged with 5'-ATT or 5'-GTT at their ends for post-labeling of both strands. DHPLC analysis is carried out using the WAVE DNA fragment analysis

15 system (Transgenomic). PCR products are injected into DNASep column, and separated under the conditions determined using WAVEMaker program (Transgenomic). The two strands of PCR products that are differentially labeled with nucleotides modified with R110 and R6G by an exchange reaction of Klenow fragment of DNA polymerase I. The reaction is stopped by adding EDTA, and unincorporated nucleotides are

20 dephosphorylated by adding calf intestinal alkaline phosphatase. SSCP followed by electrophoresis is performed in a capillary using an ABI Prism 310 Genetic Analyzer. Genescan softwares (P-E Biosystems). DNA of individuals including those who showed different genotypes on DHPLC or SSCP are subjected for direct sequencing using big-dye terminator chemistry, on ABI Prism 310 sequencer. Multiple sequence trace files

25 obtained from ABI Prism 310 are processed and aligned by Phred/Phrap and viewed using Consed viewer. SNPs are identified by PolyPhred software and visual inspection. Trace chromatogram data of EST sequences in Unigene are processed with PHRED. To identify likely SNPs, single base mismatches are reported from multiple sequence alignments produced by the programs PHRAP, BRO and POA for each Unigene cluster.

30 BRO corrected possible misreported EST orientations, while POA identified and analyzed non-linear alignment structures indicative of gene mixing/chimeras that might

produce spurious SNPs. Bayesian inference is used to weigh evidence for true polymorphism versus sequencing error, misalignment or ambiguity, misclustering or chimeric EST sequences, assessing data such as raw chromatogram height, sharpness, overlap and spacing; sequencing error rates; context-sensitivity; cDNA library origin, etc.

5 In method identified as MARSHFIELD(Method-B), overlapping human DNA sequences which contained putative insertion/deletion polymorphisms are identified through searches of public databases. PCR primers which flanked each polymorphic site are selected from the consensus sequences. Primers are used to amplify individual or pooled human genomic DNA. Resulting PCR products are resolved on a denaturing
10 polyacrylamide gel and a PhosphorImager is used to estimate allele frequencies from DNA pools.

6. Linkage Disequilibrium

Polymorphisms in linkage disequilibrium with the polymorphism at 3972 of the
15 *ABCC2* gene locus may also be used with the methods of the present invention. "Linkage disequilibrium" ("LD" as used herein, though also referred to as "LED" in the art) refers to a situation where a particular combination of alleles (*i.e.*, a variant form of a given gene) or polymorphisms at two loci appears more frequently than would be expected by chance. "Significant" as used in respect to linkage disequilibrium, as determined by one
20 of skill in the art, is contemplated to be a statistical p or α value that may be 0.25 or 0.1 and may be 0.1, 0.05, 0.001, 0.00001 or less. The relationship between *ABCC2* haplotypes and the AUC of *ABCC2* substrates may be used to correlate the genotype (*i.e.*, the genetic make up of an organism) to a phenotype (*i.e.*, the physical traits displayed by an organism or cell). "Haplotype" is used according to its plain and
25 ordinary meaning to one skilled in the art. It refers to a collective genotype of two or more alleles or polymorphisms along one of the homologous chromosomes.

A common haplotype with the 3972 variant includes two promoter variants (-1549A>G and -1019A>G) and a 5'UTR variant (-24C>T). This is found at a frequency of 17.3% in Caucasian, 4.3% in African-Americans, and 10.3% in Asian populations.
30 The 3972 variant is found alone at a frequency of 5.2% in Caucasians and 4.6% in

African-Americans. A haplotype including the 3972 variant and the -1549 and -1019 promoter variants has a frequency of 9.2% in Caucasians, and 3.7% in African-Americans. Another haplotype with the 3972 variant includes the -1549A>G promoter variant and an intronic variant in intron 13 (+27C>G). This haplotype is found at a frequency of 4.8% in African-Americans.

V. FORMULATIONS AND DOSAGES

Irinotecan is also known as CPT-11 and it is commercially available as CAMPTOSAR®. CAMPTOSAR® is supplied as a sterile solution in two single-dose sizes: 2-mL vials containing 40 mg irinotecan hydrochloride and 5-mL vials containing 100 mg irinotecan hydrochloride. Irinotecan hydrochloride is a semisynthetic derivative of camptothecin, which is an alkaloid extract from plants including *Camptotheca acuminata*.

CAMPTOSAR® Injection can be administered as a monotherapy, but in some instances is indicated as one agent of a first-line therapy to treat colon or rectal cancer. It has been used in combination with 5-fluorouracil (5-FU) and leucovorin. In some cases, this combination treatment is indicated for patient with recurrent or progressed cancer, after they have undergone a fluorouracil-based therapy.

It can be administered by intravenous infusion. Dosages of CAMPTOSAR® include 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 300, 305, 310, 315, 320, 325, 330, 335, 340, 345, 350, 355, 360, 365, 370, 375, 380, 385, 390, 400 or more mg/m² on day 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 26, 37 or on a weekly regimen, such as every 1, 2, 3, 4 weeks or more for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more consecutive or non-consecutive weeks. It is contemplated that dosages can be adjusted to be less than or more than the concentrations discussed above or less frequently or more frequently than the timing discussed above. It is contemplated treatment cycles may be repeated and that there may be a respite between cycles. One of ordinary skill in the art

is familiar with dosages regimens. In one example of a typical regimen for single-agent CAMPTOSAR® treatment, a patient is provided 125 mg/m² IV over 90 minutes on day 1, 8, 15, 22, then a two week rest before the cycle may be resumed. The overall amount of the drug administered to the patient in a single regimen or for the treatment overall may be increased or decreased by about, by at least about, or by at most about 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000 mg/m² or any ranges derivable therein.

The dosages of other ABCC2 drug substrates (drugs are included in Table 1) that are administered to patients is well known to those of skill in the art. These dosages may be reduced or increased relative to a dosage that would have been administered in the absence of genotyping. It is specifically contemplated that the dosages of any of those drugs may be similarly altered or modified based on genotypic analysis described herein.

V. KITS

Any of the compositions described herein may be comprised in a kit. In a non-limiting example, reagents for determining the genotype of one or both *ABCC2* genes are included in a kit. The kit may further include individual nucleic acids that can be used to amplify and/or detect particular nucleic acid sequences of the *ABCC2* gene. It may also include one or more buffers, such as a DNA isolation buffers, an amplification buffer or a hybridization buffer. The kit may also contain compounds and reagents to prepare DNA templates and isolate DNA from a sample. The kit may also include various labeling reagents and compounds.

The components of the kits may be packaged either in aqueous media or in lyophilized form. The container means of the kits will generally include at least one vial, test tube, flask, bottle, syringe or other container means, into which a component may be placed, and preferably, suitably aliquoted. Where there is more than one component in the kit, the kit also will generally contain a second, third or other additional container into which the additional components may be separately placed. However, various combinations of components may be comprised in a vial. The kits of the present

invention also will typically include a means for containing the nucleic acids, and any other reagent containers in close confinement for commercial sale. Such containers may include injection or blow-molded plastic containers into which the desired vials are retained.

5 When the components of the kit are provided in one and/or more liquid solutions, the liquid solution is an aqueous solution, with a sterile aqueous solution being particularly preferred. However, the components of the kit may be provided as dried powder(s). When reagents and/or components are provided as a dry powder, the powder can be reconstituted by the addition of a suitable solvent. It is envisioned that the solvent
10 may also be provided in another container means.

A kit will also include instructions for employing the kit components as well the use of any other reagent not included in the kit. Instructions may include variations that can be implemented.

It is contemplated that such reagents are embodiments of kits of the invention.
15 Such kits, however, are not limited to the particular items identified above and may include any reagent used directly or indirectly in the detection of polymorphisms in the *ABCC2* gene, particularly the 3972C>T polymorphism. Kits include, in some embodiments, nucleic acids capable of amplifying or of probing for a polymorphism in the *ABCC2* gene.

20

EXAMPLES

The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor
25 to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

30

EXAMPLE 1

CORRELATION OF THE 3972C>T VARIANT OF *ABCC2* WITH IRINOTECAN PHARMACOKINETICS

5 Sixty-four adults (48 Caucasians, 10 African-Americans, 4 Hispanics, and 2 others) with refractory solid tumors took part in the pharmacogenetic study. Genotyping of common variants ($q > 0.10$ in individuals of African and Caucasian origin) was performed for the following genes (number of variants in parenthesis): CES-2 (n=2), ABCC1 (n=7), ABCC2 (n=6), ABCB1 (n=8), CYP3A4*1B (n=1), CYP3A5*3 (n=1), 10 UGT1A9 (n=1), and HNF-1 α (n=1) (Table 2).

Gene	Location	Position
CES-2	16q22.1	-363C>G, 5'UTR
CES-2	16q22.1	1361G>A, intron 1
ABCC1	16p13.1	1062T>C, synonymous
ABCC1	16p13.1	8A>G, intron 9
ABCC1	16p13.1	-48C>, intron 11
ABCC1	16p13.1	1684T>C, synonymous
ABCC1	16p13.1	-30C>G, intron 18
ABCC1	16p13.1	4002G>A, synonymous
ABCC1	16p13.1	18A>G, intron 30
ABCC2	10q24	-1549A>G, promoter
ABCC2	10q24	-1019A>G, promoter
ABCC2	10q24	-24C>T, 5'UTR
ABCC2	10q24	1249G>A, nonsynonymous, Val417Ile
ABCC2	10q24	-34T>C, intron 26
ABCC2	10q24	3972C>T, synonymous
ABCB1	7q21.1	-129T>C, 5'UTR
ABCB1	7q21.1	-25G>T, intron 4
ABCB1	7q21.1	-44A>G, intron 9
ABCB1	7q21.1	1236C>T, synonymous
ABCB1	7q21.1	24C>T, intron 13
ABCB1	7q21.1	+38A>G, intron 14
ABCB1	7q21.1	2677G>T/A, nonsynonymous, Ala893Ser/Thr
ABCB1	7q21.1	3435C>T, synonymous
CYP3A4*1B	7q21.1	-392A>G, promoter
CYP3A5*3	7q21.1	22893G>A
UGT1A9	2q37	-11810T/9T, exon 1, AF297093
HNF1 α	12q24.2	79A>C, nonsynonymous I27L, exon 1, NM_000545.3

Table 2. Genetic variants typed in this study.

Irinotecan, SN-38, SN-38G, and APC AUCs were measured using noncompartmental analysis (WinNonlin) in the 64 patients in the study after a 350 mg/m² IV dose of irinotecan. AUC ratios of SN-38/ irinotecan, APC/ irinotecan, and SN-38G/SN-38 were also calculated. After visual inspection of the graphical plots of AUC and ratios stratified by genotype, t test analysis was applied to the data showing the possible presence of an inter-genotype difference in irinotecan pharmacokinetics.

The synonymous 3972C>T (exon 28) in ABCC2 was correlated with irinotecan AUC (p=0.02) (FIG. 1), APC AUC (p=<0.0001) (FIG. 1), and SN-38G AUC (p<0.001) (FIG. 2), with the TT patients showing higher AUC values compared to CT and CC patients. Higher values of AUC ratios in the TT patients compared to CT and CC patients were also observed in relation to APC/irinotecan (p=<0.0001) and SN-38G/SN-38 (p<0.001). For SN-38 and SN-38G AUCs, the correlation with 3972C>T was analyzed in patients with 6/6 and 6/7 UGT1A1 genotype (n=54) to avoid confounding effects of 7/7 genotypes. No significant correlation was observed between SN-38 AUC and 3972C>T (p=0.9) (FIG. 2). The frequency of CC, CT, and TT genotypes in the sample population was 0.44, 0.44, and 0.13, respectively. Other gene variants showed either no or borderline statistical significance in the anova test.

20

* * * *

All of the compositions and/or methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the compositions and/or methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents that are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and

30

modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

REFERENCES

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein
5 by reference.

U.S. Patent 4,582,788

U.S. Patent 4,659,774

U.S. Patent 4,683,194

10 U.S. Patent 4,683,195

U.S. Patent 4,683,202

U.S. Patent 4,683,202

U.S. Patent 4,683,202

U.S. Patent 4,800,159

15 U.S. Patent 4,816,571

U.S. Patent 4,883,750

U.S. Patent 4,946,773

U.S. Patent 4,959,463

U.S. Patent 5,141,813

20 U.S. Patent 5,264,566

U.S. Patent 5,279,721

U.S. Patent 5,428,148

U.S. Patent 5,554,744

U.S. Patent 5,574,146

25 U.S. Patent 5,602,244

U.S. Patent 5,645,897

U.S. Patent 5,705,629

U.S. Patent 5,840,873

U.S. Patent 5,843,640

30 U.S. Patent 5,843,650

U.S. Patent 5,843,651

U.S. Patent 5,843,663
U.S. Patent 5,846,708
U.S. Patent 5,846,709
U.S. Patent 5,846,717
5 U.S. Patent 5,846,726
U.S. Patent 5,846,729
U.S. Patent 5,846,783
U.S. Patent 5,849,481
U.S. Patent 5,849,483
10 U.S. Patent 5,849,486
U.S. Patent 5,849,487
U.S. Patent 5,849,497
U.S. Patent 5,849,546
U.S. Patent 5,849,547
15 U.S. Patent 5,851,770
U.S. Patent 5,851,772
U.S. Patent 5,853,990
U.S. Patent 5,853,992
U.S. Patent 5,853,993
20 U.S. Patent 5,856,092
U.S. Patent 5,858,652
U.S. Patent 5,861,244
U.S. Patent 5,863,732
U.S. Patent 5,863,753
25 U.S. Patent 5,866,331
U.S. Patent 5,866,337
U.S. Patent 5,866,366
U.S. Patent 5,900,481
U.S. Patent 5,905,024
30 U.S. Patent 5,910,407
U.S. Patent 5,912,124

- U.S. Patent 5,912,145
U.S. Patent 5,912,148
U.S. Patent 5,916,776
U.S. Patent 5,916,779
5 U.S. Patent 5,919,626
U.S. Patent 5,919,630
U.S. Patent 5,922,574
U.S. Patent 5,925,517
U.S. Patent 5,925,525
10 U.S. Patent 5,928,862
U.S. Patent 5,928,869
U.S. Patent 5,928,870
U.S. Patent 5,928,905
U.S. Patent 5,928,906
15 U.S. Patent 5,929,227
U.S. Patent 5,932,413
U.S. Patent 5,932,451
U.S. Patent 5,935,791
U.S. Patent 5,935,825
20 U.S. Patent 5,939,291
U.S. Patent 5,942,391
U.S. Patent 5,952,174
U.S. Patent. 4,656,127
U.S. Patent 4,682,195
25
Araki *et al.*, *Jpn. J. Cancer Res.*, 84:697-702, 1993.
Ausubel *et al.*, In: *Current Protocols in Molecular Biology*, Green Pub. Assoc., Inc., and
John Wiley & Sons, Inc., NY, (I):2.10.3, 1989.
Borst *et al.*, *Biochim. Biophys. Acta*, 1461(2):347-357, 1999
30 Borst *et al.* *J. Natl. Cancer Inst.*, 92:1295-1302, 2000.
European Application 329 822

- European Application 320 308
- European Patent 266,032
- European Patent 258,017
- European Patent 50,424
- 5 European Patent 201,184
- European Patent 237,362
- European Patent 84,796
- French Patent 2,650,840
- Froehler *et al.*, *Nucleic Acids Res.*, 14(13):5399-5407, 1986.
- 10 Frohman, In: *PCR Protocols: A Guide To Methods And Applications*, Academic Press, NY, 1990.
- Fuchs *et al.*, *J. Clin. Oncol.*, 21(5):807-814, 1993.
- Great Britain Patent 2 202 328
- Gupta *et al.*, *Cancer Res.*, 54:3723-3725, 1994.
- 15 Gupta *et al.*, *J. Clin. Oncol.*, 15:1502-1510, 1997.
- Halushka *et al.*, *Nat. Genet.*, 22(3):239-247, 1999.
- Inazuka *et al.*, *Genome Res*, 7(11):1094-1103, 1997.
- Innis *et al.*, *Proc Natl Acad Sci U S A.* 85(24):9436-9440, 1988.
- Iyer *et al.*, *J. Clin. Invest.*, 101:847-854, 1998.
- 20 Johnson *et al.*, *Nat. Genet.*, 29(2):233-237, 2001.
- Kaneda *et al.*, *Cancer Res.*, 50:1715-1720, 1990.
- Ke and Cardon, *Bioinformatics*, 19(2):287-288, 2003.
- Komher, *et al.*, *Nucl. Acids. Res.* 17:7779-7784, 1989.
- Kuppuswamy, *et al.*, *Proc. Natl. Acad. Sci. USA*, 88:1143-1147, 1991.
- 25 Kwok *et al.*, *Proc. Nat. Acad. Sci. USA*, 86:1173, 1989.
- Kwok, *Annu Rev Genomics Hum Genet.*, 2:235-58, 2001.
- Kwok and Chen, *Curr Issues Mol. Biol.*, Apr;5(2):43-60, 2003.
- Kwok *et al.*, *Genomics*, 23(1):138-144, 1994.
- Landegren, *et al.*, *Science*, 241:1077-1080, 1988.
- 30 Maxam, *et al.*, *Proc. Natl. Acad. Sci. USA*, 74:560, 1977.
- Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273, 1986.

- Negoro *et al.*, *J. Natl. Cancer Inst.*, 83(16):1164-1168, 1991.
- Nickerson *et al.*, *Proc. Natl. Acad. Sci. USA*, 87:8923-8927, 1990.
- Nyren *et al.*, *Anal. Biochem.* 208:171-175, 1993.
- Ohara *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:5673-5677, 1989.
- 5 PCT Application PCT/US87/00880
- PCT Application PCT/US89/01025
- PCT Application WO 88/10315
- PCT Application WO 89/06700
- PCT Application WO91/02087
- 10 PCT Application WO92/15712
- Prezant *et al.*, *Hum. Mutat.* 1:159-164, 1992.
- Rothenberg *et al.*, *J. Clin. Oncol.*, 11(11):2194-21204, 1993.
- Sambrook *et al.*, *In: Molecular Cloning*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
- 15 Sanger, *et al.*, *J. Molec. Biol.*, 94:441, 1975.
- Sokolov, *Nucl. Acids Res.* 18:3671, 1990.
- Sparreboom *et al.*, *Drug Resist. Update.*, 3:357-363, 2000.
- Suzuki *et al.*, *Semin. Liver. Dis.* 18:359-376, 1998.
- Suzuki *et al.*, Transporters for bile acids and organic anions, in: G.L. Amidon, W. Sadec
 20 (eds.), *Membrane Transporters as Drug Targets*, Kluwer Academic/Plenum
 Publishers, New York, 1999.
- Syvanen *et al.*, *Genomics* 8:684-692, 1990.
- Taillon-Miller *et al.*, *Genome Res*, 8(7):748-754, 1998.
- Ugozzoll *et al.*, *GATA* 9:107-112, 1992.
- 25 Vanhoefer *et al.*, *J. Clin. Oncol.*, 19(5):1501-18, 2001.
- Walker *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:392-396, 1992.

WHAT IS CLAIMED IS:

1. A method for predicting the level of ABCC2 activity in a patient comprising:
 - a) determining the sequence at position 3972 in one or both alleles of the
5 *ABCC2* gene of the patient, wherein a C at position 3972 on one or both
alleles is indicative of a normal level of ABCC2 activity.
2. The method of claim 2, wherein the sequence at position 3972 is determined for
both alleles of the *ABCC2* gene.
10
3. The method of claim 2, wherein a T at position 3972 on both alleles of the
ABCC2 gene is indicative of a lower than normal level of ABCC2 activity.
4. The method of claim 1, further comprising obtaining a sample from the patient
15 and using the sample to determine the sequence at position 3972.
5. The method of claim 4, wherein determining the sequence at position 3972 is
performed by a hybridization assay.
- 20 6. The method of claim 4, wherein determining the sequence at position 3972 is
performed by an allele specific amplification assay.
7. The method of claim 4, wherein determining the sequence at position 3972 is
performed by a sequencing or microsequencing assay.
25
8. The method of claim 4, wherein the sample comprises buccal cells, mononuclear
cells, or cancer cells.

9. The method of claim 1, wherein the sequence at position 3972 is determined by evaluating the sequence of a position in linkage disequilibrium with the sequence at position 3972.
- 5
10. The method of claim 9, wherein the position in linkage disequilibrium with the sequence at position 3972 is selected from the group consisting of positions -1549, -1019, -24, and +27.
- 10 11. The method of claim 9, wherein the sequence at position 3972 is determined by evaluating the sequence of more than one position in linkage disequilibrium with the sequence at position 3972.
- 15 12. The method of claim 1, further comprising administering an ABCC2 substrate to the patient.
13. The method of claim 1, further comprising analyzing a clearance rate for an ABCC2 substrate.
- 20 14. The method of claim 13, wherein the substrate is selected from the group consisting of irinotecan, APC, and SN-38G.
15. A method for determining dosage of an ABCC2 substrate for a patient comprising:
- 25 a) determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 on one or both

alleles indicates a higher dosage of the substrate than is indicated for a patient with a T at position 3972 in both alleles of the *ABCC2* gene.

5 16. The method of claim 15, further comprising obtaining a sample from the patient and using the sample to determine the sequence at position 3972.

17. The method of claim 16, wherein determining the sequence at position 3972 is performed by a hybridization assay.

10 18. The method of claim 16, wherein determining the sequence at position 3972 is performed by an allele specific amplification assay.

19. The method of claim 16, wherein determining the sequence at position 3972 is performed by a sequencing or microsequencing assay.

15

20. The method of claim 16, wherein the sample comprises buccal cells, mononuclear cells, or cancer cells.

20 21. The method of claim 15, wherein the sequence at position 3972 is determined by evaluating the sequence of a position in linkage disequilibrium with a sequence at position 3972.

25 22. The method of claim 15, further comprising prescribing a dosage of the substrate based on determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene.

23. A method for predicting tumor response to an anticancer agent that is an *ABCC2* substrate in a cancer patient comprising:

- 5 a) determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 on one or both alleles is indicative of a lower probability of an antitumor response to the anticancer agent.

24. The method of claim 23, wherein the sequence at position 3972 is determined for both alleles of the *ABCC2* gene.

10

25. The method of claim 24, wherein a T at position 3972 on both alleles of the *ABCC2* gene is indicative of a higher probability of an antitumor response to the anticancer agent.

15 26. The method of claim 23, further comprising obtaining a sample from the patient and using the sample to determine the sequence at position 3972.

27. The method of claim 26, wherein determining the sequence at position 3972 is performed by a hybridization assay.

20

28. The method of claim 26, wherein determining the sequence at position 3972 is performed by an allele specific amplification assay.

25 29. The method of claim 26, wherein determining the sequence at position 3972 is performed by a sequencing or microsequencing assay.

30. The method of claim 26, wherein the sample comprises buccal cells, mononuclear cells, or cancer cells.
- 5 31. The method of claim 23, wherein the sequence at position 3972 is determined by evaluating the sequence of a position in linkage disequilibrium with a sequence at position 3972.
32. The method of claim 23, further comprising administering the anticancer agent to the patient.
- 10 33. The method of claim 32, further comprising administering to the patient a second anticancer agent that is not an *ABCC2* substrate.
- 15 34. The method of claim 32, further comprising prescribing a dosage of the anticancer agent based on determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene.
- 20 35. A method for determining dosage of irinotecan for a patient comprising:
a) determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 on one or both alleles indicates a higher dosage of irinotecan than is indicated for a patient with a T at position 3972 in both alleles of the *ABCC2* gene.
- 25 36. The method of claim 35, further comprising obtaining a sample from the patient and using the sample to determine the sequence at position 3972.

37. The method of claim 36, wherein determining the sequence at position 3972 is performed by a hybridization assay.
38. The method of claim 36, wherein determining the sequence at position 3972 is performed by an allele specific amplification assay.
39. The method of claim 36, wherein determining the sequence at position 3972 is performed by a sequencing or microsequencing assay.
40. The method of claim 36, wherein the sample comprises buccal cells, mononuclear cells, or cancer cells.
41. The method of claim 35, wherein the sequence at position 3972 is determined by evaluating the sequence of a position in linkage disequilibrium with a sequence at position 3972.
42. The method of claim 35, further comprising prescribing a dosage of irinotecan based on determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene.
43. A method for predicting tumor response to irinotecan in a cancer patient comprising
- a) determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 in one or both alleles is indicative of a lower probability of an antitumor response to irinotecan.

44. The method of claim 43, wherein the sequence at position 3972 is determined for both alleles of the *ABCC2* gene.
45. The method of claim 44, wherein a T at position 3972 on both alleles of the *ABCC2* gene is indicative of a higher probability of an antitumor response to irinotecan.
46. The method of claim 43, further comprising obtaining a sample from the patient and using the sample to determine the sequence at position 3972.
47. The method of claim 46, wherein determining the sequence at position 3972 is performed by a hybridization assay.
48. The method of claim 46, wherein determining the sequence at position 3972 is performed by an allele specific amplification assay.
49. The method of claim 46, wherein determining the sequence at position 3972 is performed by a sequencing or microsequencing assay.
50. The method of claim 46, wherein the sample comprises buccal cells, mononuclear cells, or cancer cells.
51. The method of claim 43, wherein the sequence at position 3972 is determined by evaluating the sequence of a position in linkage disequilibrium with a sequence at position 3972.
52. The method of claim 43, further comprising administering irinotecan to the patient.

53. The method of claim 52, further comprising prescribing a dosage of irinotecan based on determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene.

5

54. A method for predicting a clearance rate for irinotecan in a patient comprising
a) determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene of the patient, wherein a C at position 3972 in one or both alleles is indicative of a normal clearance rate for irinotecan.

10

55. The method of claim 54, wherein the sequence at position 3972 is determined for both alleles of the *ABCC2* gene.

56. The method of claim 55, wherein a T at position 3972 on both alleles of the *ABCC2* gene is indicative of a lower than normal clearance rate for irinotecan.

15

57. The method of claim 54, further comprising obtaining a sample from the patient and using the sample to determine the sequence at position 3972.

58. The method of claim 57, wherein determining the sequence at position 3972 is performed by a hybridization assay.

20

59. The method of claim 57, wherein determining the sequence at position 3972 is performed by an allele specific amplification assay.

25

60. The method of claim 57, wherein determining the sequence at position 3972 is performed by a sequencing or microsequencing assay.

61. The method of claim 57, wherein the sample comprises buccal cells, mononuclear cells, or cancer cells.
- 5 62. The method of claim 54, wherein the sequence at position 3972 is determined by evaluating the sequence of a position in linkage disequilibrium with a sequence at position 3972.
- 10 63. The method of claim 54, further comprising administering irinotecan to the patient.
64. The method of claim 62, further comprising prescribing a dosage of irinotecan based on determining the sequence at position 3972 in one or both alleles of the *ABCC2* gene.
- 15 65. A kit for predicting a clearance rate for an *ABCC2* substrate in a patient comprising a nucleic acid for determining the sequence at position 3972 in an *ABCC2* gene.
- 20 66. The kit of claim 65, wherein the nucleic acid is a primer for amplifying the sequence at position 3972 in the *ABCC2* gene.
67. The kit of claim 65, wherein the nucleic acid is a specific hybridization probe for detecting the sequence at position 3972 in the *ABCC2* gene.
- 25 68. The kit of claim 67, wherein the specific hybridization probe is comprised in an oligonucleotide array or microarray.

ABSTRACT

The present invention is directed to methods and compositions for determining the presence or absence of polymorphisms within an ABCC2 gene and correlating these polymorphisms with activity levels of ABCC2 and making evaluations regarding the effect on ABCC2 substrates, particularly those substrates that are drugs. In addition, there are methods and compositions of evaluating the risk of an individual for developing toxicity or adverse event(s) to an ABCC2 substrate. In some embodiments, the invention concerns methods and compositions for determining the presence or absence of *ABCC2* 3972C>T variant and predicting or anticipating the level of activity of ABCC2 and determining dosages of an ABCC2 drug substrate, such as irinotecan, in a patient. Such methods and compositions can be used to evaluate whether irinotecan-based therapy, or therapy involving other ABCC2 substrates, may pose toxicity problems if given to a particular patient or predicting their efficacy. Alterations in suggested therapy may ensue based on genotyping results.

SEQUENCE LISTING

<110> RATAIN, MARK J.
INNOCENTI, FEDERICO
KROETZ, DEANNA L.
UNDEVIA, SAMIR

<120> METHODS AND COMPOSITIONS RELATING TO THE PHARMACOGENETICS
OF ABCC2 GENE VARIANTS

<130> ARCD:405USP1

<140> UNKNOWN

<141> 2004-03-05

<160> 3

<170> PatentIn Ver. 2.1

<210> 1

<211> 4868

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (38)..(4675)

<400> 1

```

gcggccgcgt ctttgttcca gacgcagtcc aggaatc atg ctg gag aag ttc tgc 55
                                         Met Leu Glu Lys Phe Cys
                                         1           5

aac tct act ttt tgg aat tcc tca ttc ctg gac agt ccg gag gca gac 103
Asn Ser Thr Phe Trp Asn Ser Ser Phe Leu Asp Ser Pro Glu Ala Asp
          10           15           20

ctg cca ctt tgt ttt gag caa act gtt ctg gtg tgg att ccc ttg ggc 151
Leu Pro Leu Cys Phe Glu Gln Thr Val Leu Val Trp Ile Pro Leu Gly
          25           30           35

ttc cta tgg ctc ctg gcc ccc tgg cag ctt ctc cac gtg tat aaa tcc 199
Phe Leu Trp Leu Leu Ala Pro Trp Gln Leu Leu His Val Tyr Lys Ser
          40           45           50

agg acc aag aga tcc tct acc acc aaa ctc tat ctt gct aag cag gta 247
Arg Thr Lys Arg Ser Ser Thr Thr Lys Leu Tyr Leu Ala Lys Gln Val
          55           60           65           70

ttc gtt ggt ttt ctt ctt att cta gca gcc ata gag ctg gcc ctt gta 295
Phe Val Gly Phe Leu Leu Ile Leu Ala Ala Ile Glu Leu Ala Leu Val
          75           80           85

ctc aca gaa gac tct gga caa gcc aca gtc cct gct gtt cga tat acc 343
Leu Thr Glu Asp Ser Gly Gln Ala Thr Val Pro Ala Val Arg Tyr Thr
          90           95           100

```

aat cca agc ctc tac cta ggc aca tgg ctc ctg gtt ttg ctg atc caa	391
Asn Pro Ser Leu Tyr Leu Gly Thr Trp Leu Leu Val Leu Leu Ile Gln	
105 110 115	
tac agc aga caa tgg tgt gta cag aaa aac tcc tgg ttc ctg tcc cta	439
Tyr Ser Arg Gln Trp Cys Val Gln Lys Asn Ser Trp Phe Leu Ser Leu	
120 125 130	
ttc tgg att ctc tcg ata ctc tgt ggc act ttc caa ttt cag act ctg	487
Phe Trp Ile Leu Ser Ile Leu Cys Gly Thr Phe Gln Phe Gln Thr Leu	
135 140 145 150	
atc cgg aca ctc tta cag ggt gac aat tct aat cta gcc tac tcc tgc	535
Ile Arg Thr Leu Leu Gln Gly Asp Asn Ser Asn Leu Ala Tyr Ser Cys	
155 160 165	
ctg ttc ttc atc tcc tac gga ttc cag atc ctg atc ctg atc ttt tca	583
Leu Phe Phe Ile Ser Tyr Gly Phe Gln Ile Leu Ile Leu Ile Phe Ser	
170 175 180	
gca ttt tca gaa aat aat gag tca tca aat aat cca tca tcc ata gct	631
Ala Phe Ser Glu Asn Asn Glu Ser Ser Asn Asn Pro Ser Ser Ile Ala	
185 190 195	
tca ttc ctg agt agc att acc tac agc tgg tat gac agc atc att ctg	679
Ser Phe Leu Ser Ser Ile Thr Tyr Ser Trp Tyr Asp Ser Ile Ile Leu	
200 205 210	
aaa ggc tac aag cgt cct ctg aca ctc gag gat gtc tgg gaa gtt gat	727
Lys Gly Tyr Lys Arg Pro Leu Thr Leu Glu Asp Val Trp Glu Val Asp	
215 220 225 230	
gaa gag atg aaa acc aag aca tta gtg agc aag ttt gaa acg cac atg	775
Glu Glu Met Lys Thr Lys Thr Leu Val Ser Lys Phe Glu Thr His Met	
235 240 245	
aag aga gag ctg cag aaa gcc agg cgg gca ctc cag aga cgg cag gag	823
Lys Arg Glu Leu Gln Lys Ala Arg Arg Ala Leu Gln Arg Arg Gln Glu	
250 255 260	
aag agc tcc cag cag aac tct gga gcc agg ctg cct ggc ttg aac aag	871
Lys Ser Ser Gln Gln Asn Ser Gly Ala Arg Leu Pro Gly Leu Asn Lys	
265 270 275	
aat cag agt caa agc caa gat gcc ctt gtc ctg gaa gat gtt gaa aag	919
Asn Gln Ser Gln Ser Gln Asp Ala Leu Val Leu Glu Asp Val Glu Lys	
280 285 290	
aaa aaa aag aag tct ggg acc aaa aaa gat gtt cca aaa tcc tgg ttg	967
Lys Lys Lys Lys Ser Gly Thr Lys Lys Asp Val Pro Lys Ser Trp Leu	
295 300 305 310	
atg aag gct ctg ttc aaa act ttc tac atg gtg ctc ctg aaa tca ttc	1015
Met Lys Ala Leu Phe Lys Thr Phe Tyr Met Val Leu Leu Lys Ser Phe	
315 320 325	

cta ctg aag cta gtg aat gac atc ttc acg ttt gtg agt cct cag ctg	1063
Leu Leu Lys Leu Val Asn Asp Ile Phe Thr Phe Val Ser Pro Gln Leu	
330 335 340	
ctg aaa ttg ctg atc tcc ttt gca agt gac cgt gac aca tat ttg tgg	1111
Leu Lys Leu Leu Ile Ser Phe Ala Ser Asp Arg Asp Thr Tyr Leu Trp	
345 350 355	
att gga tat ctg tgt gca atc ctg tta ttc act gcg gct ctg att cag	1159
Ile Gly Tyr Leu Cys Ala Ile Leu Leu Phe Thr Ala Ala Leu Ile Gln	
360 365 370	
tct ttc tgc ctt cag tgt tat ttc caa ctg tgc ttc aag ctg ggt gta	1207
Ser Phe Cys Leu Gln Cys Tyr Phe Gln Leu Cys Phe Lys Leu Gly Val	
375 380 385 390	
aaa gta cgg aca gct atc atg gct tct gta tat aag aag gca ttg acc	1255
Lys Val Arg Thr Ala Ile Met Ala Ser Val Tyr Lys Lys Ala Leu Thr	
395 400 405	
cta tcc aac ttg gcc agg aag gag tac acc gtt gga gaa aca gtg aac	1303
Leu Ser Asn Leu Ala Arg Lys Glu Tyr Thr Val Gly Glu Thr Val Asn	
410 415 420	
ctg atg tct gtg gat gcc cag aag ctg atg gat gtg acc aac ttc atg	1351
Leu Met Ser Val Asp Ala Gln Lys Leu Met Asp Val Thr Asn Phe Met	
425 430 435	
cac atg ctg tgg tca agt gtt cta cag att gtc tta tct atc ttc ttc	1399
His Met Leu Trp Ser Ser Val Leu Gln Ile Val Leu Ser Ile Phe Phe	
440 445 450	
cta tgg aga gag ttg gga ccc tca gtc tta gca ggt gtt ggg gtg atg	1447
Leu Trp Arg Glu Leu Gly Pro Ser Val Leu Ala Gly Val Gly Val Met	
455 460 465 470	
gtg ctt gta atc cca att aat gcg ata ctg tcc acc aag agt aag acc	1495
Val Leu Val Ile Pro Ile Asn Ala Ile Leu Ser Thr Lys Ser Lys Thr	
475 480 485	
att cag gtc aaa aat atg aag aat aaa gac aaa cgt tta aag atc atg	1543
Ile Gln Val Lys Asn Met Lys Asn Lys Asp Lys Arg Leu Lys Ile Met	
490 495 500	
aat gag att ctt agt gga atc aag atc ctg aaa tat ttt gcc tgg gaa	1591
Asn Glu Ile Leu Ser Gly Ile Lys Ile Leu Lys Tyr Phe Ala Trp Glu	
505 510 515	
cct tca ttc aga gac caa gta caa aac ctg cgg aag aaa gag ctg aag	1639
Pro Ser Phe Arg Asp Gln Val Gln Asn Leu Arg Lys Lys Glu Leu Lys	
520 525 530	
aac ctg ctg gcc ttt agt caa cta cag tgt gta gta ata ttc gtc ttc	1687
Asn Leu Leu Ala Phe Ser Gln Leu Gln Cys Val Val Ile Phe Val Phe	
535 540 545 550	
cag tta act cca gtc ctg gta tct gtg gtc aca ttt tct gtt tat gtc	1735

Gln	Leu	Thr	Pro	Val	Leu	Val	Ser	Val	Val	Thr	Phe	Ser	Val	Tyr	Val		
				555					560					565			
ctg	gtg	gat	agc	aac	aat	att	ttg	gat	gca	caa	aag	gcc	ttc	acc	tcc	1783	
Leu	Val	Asp	Ser	Asn	Asn	Ile	Leu	Asp	Ala	Gln	Lys	Ala	Phe	Thr	Ser		
			570					575					580				
att	acc	ctc	ttc	aat	atc	ctg	cgc	ttt	ccc	ctg	agc	atg	ctt	ccc	atg	1831	
Ile	Thr	Leu	Phe	Asn	Ile	Leu	Arg	Phe	Pro	Leu	Ser	Met	Leu	Pro	Met		
		585					590					595					
atg	atc	tcc	tcc	atg	ctc	cag	gcc	agt	gtt	tcc	aca	gag	cgg	cta	gag	1879	
Met	Ile	Ser	Ser	Met	Leu	Gln	Ala	Ser	Val	Ser	Thr	Glu	Arg	Leu	Glu		
	600					605					610						
aag	tac	ttg	gga	ggg	gat	gac	ttg	gac	aca	tct	gcc	att	cga	cat	gac	1927	
Lys	Tyr	Leu	Gly	Gly	Asp	Asp	Leu	Asp	Thr	Ser	Ala	Ile	Arg	His	Asp		
615					620					625					630		
tgc	aat	ttt	gac	aaa	gcc	atg	cag	ttt	tct	gag	gcc	tcc	ttt	acc	tgg	1975	
Cys	Asn	Phe	Asp	Lys	Ala	Met	Gln	Phe	Ser	Glu	Ala	Ser	Phe	Thr	Trp		
			635						640					645			
gaa	cat	gat	tcg	gaa	gcc	aca	gtc	cga	gat	gtg	aac	ctg	gac	att	atg	2023	
Glu	His	Asp	Ser	Glu	Ala	Thr	Val	Arg	Asp	Val	Asn	Leu	Asp	Ile	Met		
			650					655					660				
gca	ggc	caa	ctt	gtg	gct	gtg	ata	ggc	cct	gtc	ggc	tct	ggg	aaa	tcc	2071	
Ala	Gly	Gln	Leu	Val	Ala	Val	Ile	Gly	Pro	Val	Gly	Ser	Gly	Lys	Ser		
		665					670					675					
tcc	ttg	ata	tca	gcc	atg	ctg	gga	gaa	atg	gaa	aat	gtc	cac	ggg	cac	2119	
Ser	Leu	Ile	Ser	Ala	Met	Leu	Gly	Glu	Met	Glu	Asn	Val	His	Gly	His		
	680					685					690						
atc	acc	atc	aag	ggc	acc	act	gcc	tat	gtc	cca	cag	cag	tcc	tgg	att	2167	
Ile	Thr	Ile	Lys	Gly	Thr	Thr	Ala	Tyr	Val	Pro	Gln	Gln	Ser	Trp	Ile		
695					700				705					710			
cag	aat	ggc	acc	ata	aag	gac	aac	atc	ctt	ttt	gga	aca	gag	ttt	aat	2215	
Gln	Asn	Gly	Thr	Ile	Lys	Asp	Asn	Ile	Leu	Phe	Gly	Thr	Glu	Phe	Asn		
			715					720					725				
gaa	aag	agg	tac	cag	caa	gta	ctg	gag	gcc	tgt	gct	ctc	ctc	cca	gac	2263	
Glu	Lys	Arg	Tyr	Gln	Gln	Val	Leu	Glu	Ala	Cys	Ala	Leu	Leu	Pro	Asp		
			730					735					740				
ttg	gaa	atg	ctg	cct	gga	gga	gat	ttg	gct	gag	att	gga	gag	aag	ggc	2311	
Leu	Glu	Met	Leu	Pro	Gly	Gly	Asp	Leu	Ala	Glu	Ile	Gly	Glu	Lys	Gly		
		745					750					755					
ata	aat	ctt	agt	ggg	ggc	cag	aag	cag	cgg	atc	agc	ctg	gcc	aga	gct	2359	
Ile	Asn	Leu	Ser	Gly	Gly	Gln	Lys	Gln	Arg	Ile	Ser	Leu	Ala	Arg	Ala		
	760					765					770						
acc	tac	caa	aat	tta	gac	atc	tat	ctt	cta	gat	gac	ccc	ctg	tct	gca	2407	
Thr	Tyr	Gln	Asn	Leu	Asp	Ile	Tyr	Leu	Leu	Asp	Asp	Pro	Leu	Ser	Ala		

775	780	785	790	
gtg gat gct cat	gta gga aaa cat	att ttt aat aag	gtc ttg ggc ccc	2455
Val Asp Ala His	Val Gly Lys His	Ile Phe Asn Lys	Val Leu Gly Pro	
	795	800	805	
aat ggc ctg ttg	aaa ggc aag act	cga ctc ttg gtt	aca cat agc atg	2503
Asn Gly Leu Leu	Lys Gly Lys Thr	Arg Leu Leu Val	Thr His Ser Met	
	810	815	820	
cac ttt ctt cct	caa gtg gat gag	att gta gtt ctg	ggg aat gga aca	2551
His Phe Leu Pro	Gln Val Asp Glu	Ile Val Val Leu	Gly Asn Gly Thr	
	825	830	835	
att gta gag aaa	gga tcc tac agt	gct ctc ctg gcc	aaa aaa gga gag	2599
Ile Val Glu Lys	Gly Ser Tyr Ser	Ala Leu Leu Ala	Lys Lys Gly Glu	
	840	845	850	
ttt gct aag aat	ctg aag aca ttt	cta aga cat aca	ggc cct gaa gag	2647
Phe Ala Lys Asn	Leu Lys Thr Phe	Leu Arg His Thr	Gly Pro Glu Glu	
	860	865	870	
gaa gcc aca gtc	cat gat ggc agt	gaa gaa gaa gac	gat gac tat ggg	2695
Glu Ala Thr Val	His Asp Gly Ser	Glu Glu Glu Asp	Asp Asp Tyr Gly	
	875	880	885	
ctg ata tcc agt	gtg gaa gag atc	ccc gaa gat gca	gcc tcc ata acc	2743
Leu Ile Ser Ser	Val Glu Glu Ile	Pro Glu Asp Ala	Ala Ser Ile Thr	
	890	895	900	
atg aga aga gag	aac agc ttt cgt	cga aca ctt agc	cgc agt tct agg	2791
Met Arg Arg Glu	Asn Ser Phe Arg	Arg Thr Leu Ser	Arg Ser Ser Arg	
	905	910	915	
tcc aat ggc agg	cat ctg aag tcc	ctg aga aac tcc	ttg aaa act cgg	2839
Ser Asn Gly Arg	His Leu Lys Ser	Leu Arg Asn Ser	Leu Lys Thr Arg	
	920	925	930	
aat gtg aat agc	ctg aag gaa gac	gaa gaa cta gtg	aaa gga caa aaa	2887
Asn Val Asn Ser	Leu Lys Glu Asp	Glu Glu Leu Val	Lys Gly Gln Lys	
	940	945	950	
cta att aag aag	gaa ttc ata gaa	act gga aag gtg	aag ttc tcc atc	2935
Leu Ile Lys Lys	Glu Phe Ile Glu	Thr Gly Lys Val	Lys Phe Ser Ile	
	955	960	965	
tac ctg gag tac	cta caa gca ata	gga ttg ttt tcg	ata ttc ttc atc	2983
Tyr Leu Glu Tyr	Leu Gln Ala Ile	Gly Leu Phe Ser	Ile Phe Phe Ile	
	970	975	980	
atc ctt gcg ttt	gtg atg aat tct	gtg gct ttt att	gga tcc aac ctc	3031
Ile Leu Ala Phe	Val Met Asn Ser	Val Ala Phe Ile	Gly Ser Asn Leu	
	985	990	995	
tgg ctc agt gct	tgg acc agt gac	tct aaa atc ttc	aat agc acc gac	3079
Trp Leu Ser Ala	Trp Thr Ser Asp	Ser Lys Ile Phe	Asn Ser Thr Asp	
	1000	1005	1010	

ggc ttt gtt ctg tcc aat gca ctc aat atc aca caa acc ctg aac tgg	3799
Gly Phe Val Leu Ser Asn Ala Leu Asn Ile Thr Gln Thr Leu Asn Trp	
1240 1245 1250	
ctg gtg agg atg aca tca gaa ata gag acc aac att gtg gct gtt gag	3847
Leu Val Arg Met Thr Ser Glu Ile Glu Thr Asn Ile Val Ala Val Glu	
1255 1260 1265 1270	
cga ata act gag tac aca aaa gtg gaa aat gag gca ccc tgg gtg act	3895
Arg Ile Thr Glu Tyr Thr Lys Val Glu Asn Glu Ala Pro Trp Val Thr	
1275 1280 1285	
gat aag agg cct ccg cca gat tgg ccc agc aaa ggc aag atc cag ttt	3943
Asp Lys Arg Pro Pro Pro Asp Trp Pro Ser Lys Gly Lys Ile Gln Phe	
1290 1295 1300	
aac aac tac caa gtg cgg tac cga cct gag ctg gat ctg gtc ctc aga	3991
Asn Asn Tyr Gln Val Arg Tyr Arg Pro Glu Leu Asp Leu Val Leu Arg	
1305 1310 1315	
ggg atc act tgt gac atc ggt agc atg gag aag att ggt gtg gtg ggc	4039
Gly Ile Thr Cys Asp Ile Gly Ser Met Glu Lys Ile Gly Val Val Gly	
1320 1325 1330	
agg aca gga gct gga aag tca tcc ctc aca aac tgc ctc ttc aga atc	4087
Arg Thr Gly Ala Gly Lys Ser Ser Leu Thr Asn Cys Leu Phe Arg Ile	
1335 1340 1345 1350	
tta gag gct gcc ggt ggt cag att atc att gat gga gta gat att gct	4135
Leu Glu Ala Ala Gly Gly Gln Ile Ile Ile Asp Gly Val Asp Ile Ala	
1355 1360 1365	
tcc att ggg ctc cac gac ctc cga gag aag ctg acc atc atc ccc cag	4183
Ser Ile Gly Leu His Asp Leu Arg Glu Lys Leu Thr Ile Ile Pro Gln	
1370 1375 1380	
gac ccc atc ctg ttc tct gga agc ctg agg atg aat ctc gac cct ttc	4231
Asp Pro Ile Leu Phe Ser Gly Ser Leu Arg Met Asn Leu Asp Pro Phe	
1385 1390 1395	
aac aac tac tca gat gag gag att tgg aag gcc ttg gag ctg gct cac	4279
Asn Asn Tyr Ser Asp Glu Glu Ile Trp Lys Ala Leu Glu Leu Ala His	
1400 1405 1410	
ctc aag tct ttt gtg gcc agc ctg caa ctt ggg tta tcc cac gaa gtg	4327
Leu Lys Ser Phe Val Ala Ser Leu Gln Leu Gly Leu Ser His Glu Val	
1415 1420 1425 1430	
aca gag gct ggt ggc aac ctg agc ata ggc cag agg cag ctg ctg tgc	4375
Thr Glu Ala Gly Gly Asn Leu Ser Ile Gly Gln Arg Gln Leu Leu Cys	
1435 1440 1445	
ctg ggc agg gct ctg ctt cgg aaa tcc aag atc ctg gtc ctg gat gag	4423
Leu Gly Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu Val Leu Asp Glu	
1450 1455 1460	
gcc act gct gcg gtg gat cta gag aca gac aac ctc att cag acg acc	4471

Ala Thr Ala Ala Val Asp Leu Glu Thr Asp Asn Leu Ile Gln Thr Thr
1465 1470 1475

atc caa aac gag ttc gcc cac tgc aca gtg atc acc atc gcc cac agg 4519
Ile Gln Asn Glu Phe Ala His Cys Thr Val Ile Thr Ile Ala His Arg
1480 1485 1490

ctg cac acc atc atg gac agt gac aag gta atg gtc cta gac aac ggg 4567
Leu His Thr Ile Met Asp Ser Asp Lys Val Met Val Leu Asp Asn Gly
1495 1500 1505 1510

aag att ata gag tgc ggc agc cct gaa gaa ctg cta caa atc cct gga 4615
Lys Ile Ile Glu Cys Gly Ser Pro Glu Glu Leu Leu Gln Ile Pro Gly
1515 1520 1525

ccc ttt tac ttt atg gct aag gaa gct ggc att gag aat gtg aac agc 4663
Pro Phe Tyr Phe Met Ala Lys Glu Ala Gly Ile Glu Asn Val Asn Ser
1530 1535 1540

aca aaa ttc tag cagaaggccc catggggttag aaaaggacta taagaataat 4715
Thr Lys Phe
1545

ttcttatttta attttatttt ttataaaata cagaatacat acaaaagtgt gtataaaatg 4775

tacgtttttaa aaaaggataa gtgaacaccc atgaacctac taccaggtt aagaaaataa 4835

atgtcaccag gtacttgaga aaccctcga ttg 4868

<210> 2
<211> 1545
<212> PRT
<213> Homo sapiens

<400> 2
Met Leu Glu Lys Phe Cys Asn Ser Thr Phe Trp Asn Ser Ser Phe Leu
1 5 10 15
Asp Ser Pro Glu Ala Asp Leu Pro Leu Cys Phe Glu Gln Thr Val Leu
20 25 30
Val Trp Ile Pro Leu Gly Phe Leu Trp Leu Leu Ala Pro Trp Gln Leu
35 40 45
Leu His Val Tyr Lys Ser Arg Thr Lys Arg Ser Ser Thr Thr Lys Leu
50 55 60
Tyr Leu Ala Lys Gln Val Phe Val Gly Phe Leu Leu Ile Leu Ala Ala
65 70 75 80
Ile Glu Leu Ala Leu Val Leu Thr Glu Asp Ser Gly Gln Ala Thr Val
85 90 95
Pro Ala Val Arg Tyr Thr Asn Pro Ser Leu Tyr Leu Gly Thr Trp Leu
100 105 110
Leu Val Leu Leu Ile Gln Tyr Ser Arg Gln Trp Cys Val Gln Lys Asn
115 120 125
Ser Trp Phe Leu Ser Leu Phe Trp Ile Leu Ser Ile Leu Cys Gly Thr
130 135 140
Phe Gln Phe Gln Thr Leu Ile Arg Thr Leu Leu Gln Gly Asp Asn Ser
145 150 155 160
Asn Leu Ala Tyr Ser Cys Leu Phe Phe Ile Ser Tyr Gly Phe Gln Ile

25391509.1

				165				170				175			
Leu	Ile	Leu	Ile	Phe	Ser	Ala	Phe	Ser	Glu	Asn	Asn	Glu	Ser	Ser	Asn
			180					185				190			
Asn	Pro	Ser	Ser	Ile	Ala	Ser	Phe	Leu	Ser	Ser	Ile	Thr	Tyr	Ser	Trp
		195					200					205			
Tyr	Asp	Ser	Ile	Ile	Leu	Lys	Gly	Tyr	Lys	Arg	Pro	Leu	Thr	Leu	Glu
	210					215					220				
Asp	Val	Trp	Glu	Val	Asp	Glu	Glu	Met	Lys	Thr	Lys	Thr	Leu	Val	Ser
	225				230					235					240
Lys	Phe	Glu	Thr	His	Met	Lys	Arg	Glu	Leu	Gln	Lys	Ala	Arg	Arg	Ala
			245					250						255	
Leu	Gln	Arg	Arg	Gln	Glu	Lys	Ser	Ser	Gln	Gln	Asn	Ser	Gly	Ala	Arg
		260						265					270		
Leu	Pro	Gly	Leu	Asn	Lys	Asn	Gln	Ser	Gln	Ser	Gln	Asp	Ala	Leu	Val
	275						280					285			
Leu	Glu	Asp	Val	Glu	Lys	Lys	Lys	Lys	Lys	Ser	Gly	Thr	Lys	Lys	Asp
	290					295					300				
Val	Pro	Lys	Ser	Trp	Leu	Met	Lys	Ala	Leu	Phe	Lys	Thr	Phe	Tyr	Met
	305				310					315					320
Val	Leu	Leu	Lys	Ser	Phe	Leu	Leu	Lys	Leu	Val	Asn	Asp	Ile	Phe	Thr
			325					330						335	
Phe	Val	Ser	Pro	Gln	Leu	Leu	Lys	Leu	Leu	Ile	Ser	Phe	Ala	Ser	Asp
		340						345				350			
Arg	Asp	Thr	Tyr	Leu	Trp	Ile	Gly	Tyr	Leu	Cys	Ala	Ile	Leu	Leu	Phe
	355						360					365			
Thr	Ala	Ala	Leu	Ile	Gln	Ser	Phe	Cys	Leu	Gln	Cys	Tyr	Phe	Gln	Leu
	370				375						380				
Cys	Phe	Lys	Leu	Gly	Val	Lys	Val	Arg	Thr	Ala	Ile	Met	Ala	Ser	Val
	385				390					395					400
Tyr	Lys	Lys	Ala	Leu	Thr	Leu	Ser	Asn	Leu	Ala	Arg	Lys	Glu	Tyr	Thr
			405					410					415		
Val	Gly	Glu	Thr	Val	Asn	Leu	Met	Ser	Val	Asp	Ala	Gln	Lys	Leu	Met
		420						425				430			
Asp	Val	Thr	Asn	Phe	Met	His	Met	Leu	Trp	Ser	Ser	Val	Leu	Gln	Ile
	435						440					445			
Val	Leu	Ser	Ile	Phe	Phe	Leu	Trp	Arg	Glu	Leu	Gly	Pro	Ser	Val	Leu
	450					455					460				
Ala	Gly	Val	Gly	Val	Met	Val	Leu	Val	Ile	Pro	Ile	Asn	Ala	Ile	Leu
	465				470					475					480
Ser	Thr	Lys	Ser	Lys	Thr	Ile	Gln	Val	Lys	Asn	Met	Lys	Asn	Lys	Asp
			485					490						495	
Lys	Arg	Leu	Lys	Ile	Met	Asn	Glu	Ile	Leu	Ser	Gly	Ile	Lys	Ile	Leu
		500						505					510		
Lys	Tyr	Phe	Ala	Trp	Glu	Pro	Ser	Phe	Arg	Asp	Gln	Val	Gln	Asn	Leu
	515						520					525			
Arg	Lys	Lys	Glu	Leu	Lys	Asn	Leu	Leu	Ala	Phe	Ser	Gln	Leu	Gln	Cys
	530					535					540				
Val	Val	Ile	Phe	Val	Phe	Gln	Leu	Thr	Pro	Val	Leu	Val	Ser	Val	Val
	545				550					555					560
Thr	Phe	Ser	Val	Tyr	Val	Leu	Val	Asp	Ser	Asn	Asn	Ile	Leu	Asp	Ala
			565					570					575		
Gln	Lys	Ala	Phe	Thr	Ser	Ile	Thr	Leu	Phe	Asn	Ile	Leu	Arg	Phe	Pro
			580					585					590		
Leu	Ser	Met	Leu	Pro	Met	Met	Ile	Ser	Ser	Met	Leu	Gln	Ala	Ser	Val
	595						600					605			
Ser	Thr	Glu	Arg	Leu	Glu	Lys	Tyr	Leu	Gly	Gly	Asp	Asp	Leu	Asp	Thr
	610					615						620			

Ser	Ala	Ile	Arg	His	Asp	Cys	Asn	Phe	Asp	Lys	Ala	Met	Gln	Phe	Ser	625	630	635	640
Glu	Ala	Ser	Phe	Thr	Trp	Glu	His	Asp	Ser	Glu	Ala	Thr	Val	Arg	Asp	645	650	655	
Val	Asn	Leu	Asp	Ile	Met	Ala	Gly	Gln	Leu	Val	Ala	Val	Ile	Gly	Pro	660	665	670	
Val	Gly	Ser	Gly	Lys	Ser	Ser	Leu	Ile	Ser	Ala	Met	Leu	Gly	Glu	Met	675	680	685	
Glu	Asn	Val	His	Gly	His	Ile	Thr	Ile	Lys	Gly	Thr	Thr	Ala	Tyr	Val	690	695	700	
Pro	Gln	Gln	Ser	Trp	Ile	Gln	Asn	Gly	Thr	Ile	Lys	Asp	Asn	Ile	Leu	705	710	715	720
Phe	Gly	Thr	Glu	Phe	Asn	Glu	Lys	Arg	Tyr	Gln	Gln	Val	Leu	Glu	Ala	725	730	735	
Cys	Ala	Leu	Leu	Pro	Asp	Leu	Glu	Met	Leu	Pro	Gly	Gly	Asp	Leu	Ala	740	745	750	
Glu	Ile	Gly	Glu	Lys	Gly	Ile	Asn	Leu	Ser	Gly	Gly	Gln	Lys	Gln	Arg	755	760	765	
Ile	Ser	Leu	Ala	Arg	Ala	Thr	Tyr	Gln	Asn	Leu	Asp	Ile	Tyr	Leu	Leu	770	775	780	
Asp	Asp	Pro	Leu	Ser	Ala	Val	Asp	Ala	His	Val	Gly	Lys	His	Ile	Phe	785	790	795	800
Asn	Lys	Val	Leu	Gly	Pro	Asn	Gly	Leu	Leu	Lys	Gly	Lys	Thr	Arg	Leu	805	810	815	
Leu	Val	Thr	His	Ser	Met	His	Phe	Leu	Pro	Gln	Val	Asp	Glu	Ile	Val	820	825	830	
Val	Leu	Gly	Asn	Gly	Thr	Ile	Val	Glu	Lys	Gly	Ser	Tyr	Ser	Ala	Leu	835	840	845	
Leu	Ala	Lys	Lys	Gly	Glu	Phe	Ala	Lys	Asn	Leu	Lys	Thr	Phe	Leu	Arg	850	855	860	
His	Thr	Gly	Pro	Glu	Glu	Glu	Ala	Thr	Val	His	Asp	Gly	Ser	Glu	Glu	865	870	875	880
Glu	Asp	Asp	Asp	Tyr	Gly	Leu	Ile	Ser	Ser	Val	Glu	Glu	Ile	Pro	Glu	885	890	895	
Asp	Ala	Ala	Ser	Ile	Thr	Met	Arg	Arg	Glu	Asn	Ser	Phe	Arg	Arg	Thr	900	905	910	
Leu	Ser	Arg	Ser	Ser	Arg	Ser	Asn	Gly	Arg	His	Leu	Lys	Ser	Leu	Arg	915	920	925	
Asn	Ser	Leu	Lys	Thr	Arg	Asn	Val	Asn	Ser	Leu	Lys	Glu	Asp	Glu	Glu	930	935	940	
Leu	Val	Lys	Gly	Gln	Lys	Leu	Ile	Lys	Lys	Glu	Phe	Ile	Glu	Thr	Gly	945	950	955	960
Lys	Val	Lys	Phe	Ser	Ile	Tyr	Leu	Glu	Tyr	Leu	Gln	Ala	Ile	Gly	Leu	965	970	975	
Phe	Ser	Ile	Phe	Phe	Ile	Ile	Leu	Ala	Phe	Val	Met	Asn	Ser	Val	Ala	980	985	990	
Phe	Ile	Gly	Ser	Asn	Leu	Trp	Leu	Ser	Ala	Trp	Thr	Ser	Asp	Ser	Lys	995	1000	1005	
Ile	Phe	Asn	Ser	Thr	Asp	Tyr	Pro	Ala	Ser	Gln	Arg	Asp	Met	Arg	Val	1010	1015	1020	
Gly	Val	Tyr	Gly	Ala	Leu	Gly	Leu	Ala	Gln	Gly	Ile	Phe	Val	Phe	Ile	1025	1030	1035	1040
Ala	His	Phe	Trp	Ser	Ala	Phe	Gly	Phe	Val	His	Ala	Ser	Asn	Ile	Leu	1045	1050	1055	
His	Lys	Gln	Leu	Leu	Asn	Asn	Ile	Leu	Arg	Ala	Pro	Met	Arg	Phe	Phe	1060	1065	1070	
Asp	Thr	Thr	Pro	Thr	Gly	Arg	Ile	Val	Asn	Arg	Phe	Ala	Gly	Asp	Ile				

1075	1080	1085
Ser Thr Val Asp Asp Thr Leu Pro Gln Ser Leu Arg Ser Trp Ile Thr		
1090	1095	1100
Cys Phe Leu Gly Ile Ile Ser Thr Leu Val Met Ile Cys Met Ala Thr		
1105	1110	1115
Pro Val Phe Thr Ile Ile Val Ile Pro Leu Gly Ile Ile Tyr Val Ser		
1125	1130	1135
Val Gln Met Phe Tyr Val Ser Thr Ser Arg Gln Leu Arg Arg Leu Asp		
1140	1145	1150
Ser Val Thr Arg Ser Pro Ile Tyr Ser His Phe Ser Glu Thr Val Ser		
1155	1160	1165
Gly Leu Pro Val Ile Arg Ala Phe Glu His Gln Gln Arg Phe Leu Lys		
1170	1175	1180
His Asn Glu Val Arg Ile Asp Thr Asn Gln Lys Cys Val Phe Ser Trp		
1185	1190	1195
Ile Thr Ser Asn Arg Trp Leu Ala Ile Arg Leu Glu Leu Val Gly Asn		
1205	1210	1215
Leu Thr Val Phe Phe Ser Ala Leu Met Met Val Ile Tyr Arg Asp Thr		
1220	1225	1230
Leu Ser Gly Asp Thr Val Gly Phe Val Leu Ser Asn Ala Leu Asn Ile		
1235	1240	1245
Thr Gln Thr Leu Asn Trp Leu Val Arg Met Thr Ser Glu Ile Glu Thr		
1250	1255	1260
Asn Ile Val Ala Val Glu Arg Ile Thr Glu Tyr Thr Lys Val Glu Asn		
1265	1270	1275
Glu Ala Pro Trp Val Thr Asp Lys Arg Pro Pro Pro Asp Trp Pro Ser		
1285	1290	1295
Lys Gly Lys Ile Gln Phe Asn Asn Tyr Gln Val Arg Tyr Arg Pro Glu		
1300	1305	1310
Leu Asp Leu Val Leu Arg Gly Ile Thr Cys Asp Ile Gly Ser Met Glu		
1315	1320	1325
Lys Ile Gly Val Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Thr		
1330	1335	1340
Asn Cys Leu Phe Arg Ile Leu Glu Ala Ala Gly Gly Gln Ile Ile Ile		
1345	1350	1355
Asp Gly Val Asp Ile Ala Ser Ile Gly Leu His Asp Leu Arg Glu Lys		
1365	1370	1375
Leu Thr Ile Ile Pro Gln Asp Pro Ile Leu Phe Ser Gly Ser Leu Arg		
1380	1385	1390
Met Asn Leu Asp Pro Phe Asn Asn Tyr Ser Asp Glu Glu Ile Trp Lys		
1395	1400	1405
Ala Leu Glu Leu Ala His Leu Lys Ser Phe Val Ala Ser Leu Gln Leu		
1410	1415	1420
Gly Leu Ser His Glu Val Thr Glu Ala Gly Gly Asn Leu Ser Ile Gly		
1425	1430	1435
Gln Arg Gln Leu Leu Cys Leu Gly Arg Ala Leu Leu Arg Lys Ser Lys		
1445	1450	1455
Ile Leu Val Leu Asp Glu Ala Thr Ala Ala Val Asp Leu Glu Thr Asp		
1460	1465	1470
Asn Leu Ile Gln Thr Thr Ile Gln Asn Glu Phe Ala His Cys Thr Val		
1475	1480	1485
Ile Thr Ile Ala His Arg Leu His Thr Ile Met Asp Ser Asp Lys Val		
1490	1495	1500
Met Val Leu Asp Asn Gly Lys Ile Ile Glu Cys Gly Ser Pro Glu Glu		
1505	1510	1515
Leu Leu Gln Ile Pro Gly Pro Phe Tyr Phe Met Ala Lys Glu Ala Gly		
1525	1530	1535

Ile Glu Asn Val Asn Ser Thr Lys Phe
1540 1545

<210> 3
<211> 184
<212> DNA
<213> Homo sapiens

<400> 3
aacttacttc tcattctgtc tccttgccag gcaccctggg tgactgataa gaggcctccg 60
ccagattggc ccagcaaagg caagatccag tttaacaact accaagtgcg gtaccgacct 120
gagctggatc tggtcctcag agggatcact tgtgacatcg gtagcatgga gaaggtaggt 180
ggag 184

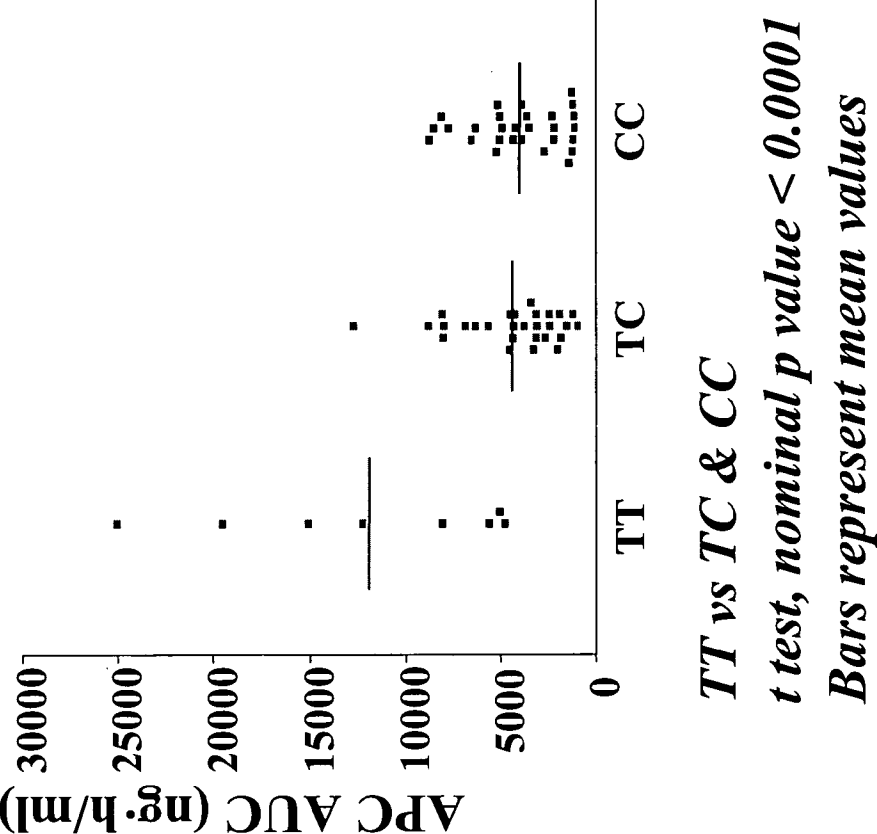
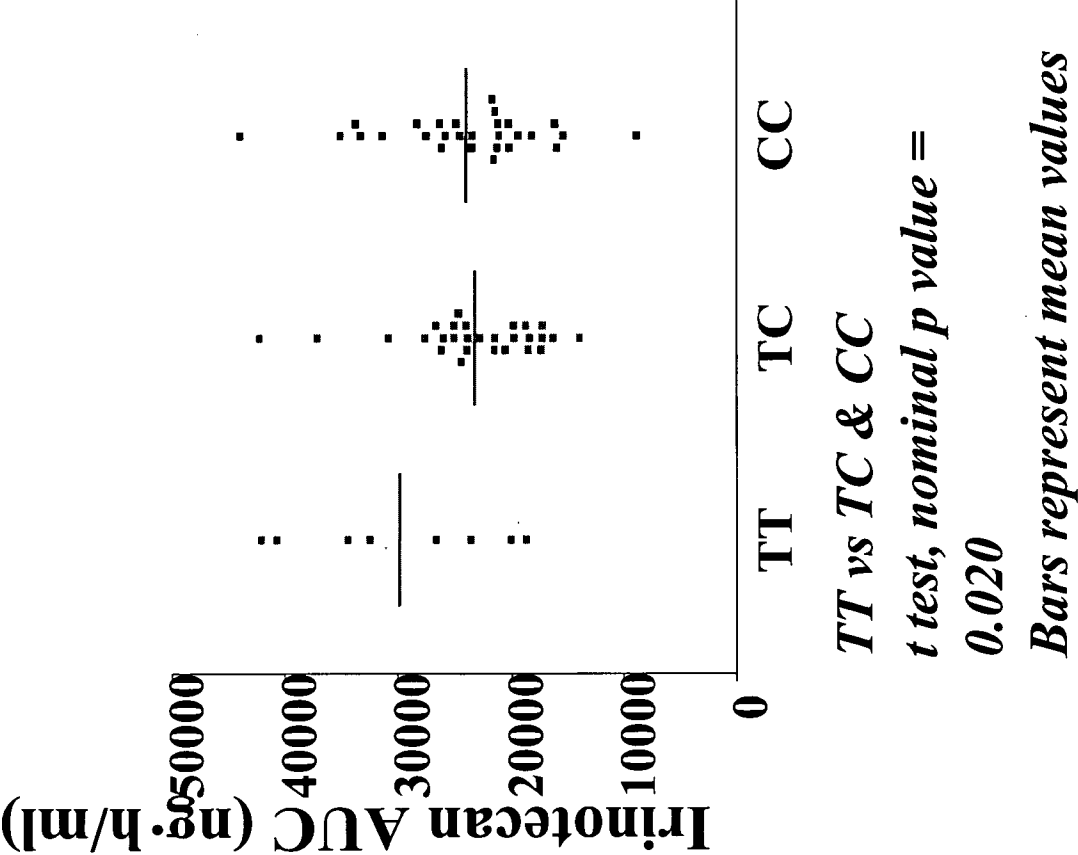


FIG. 1

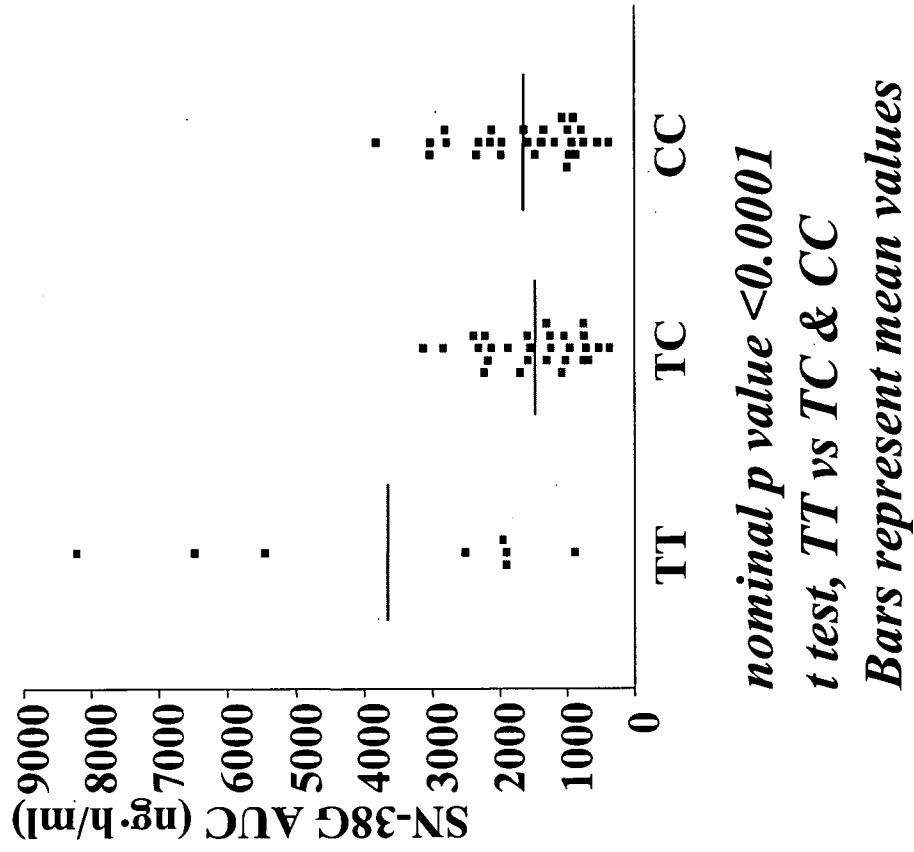
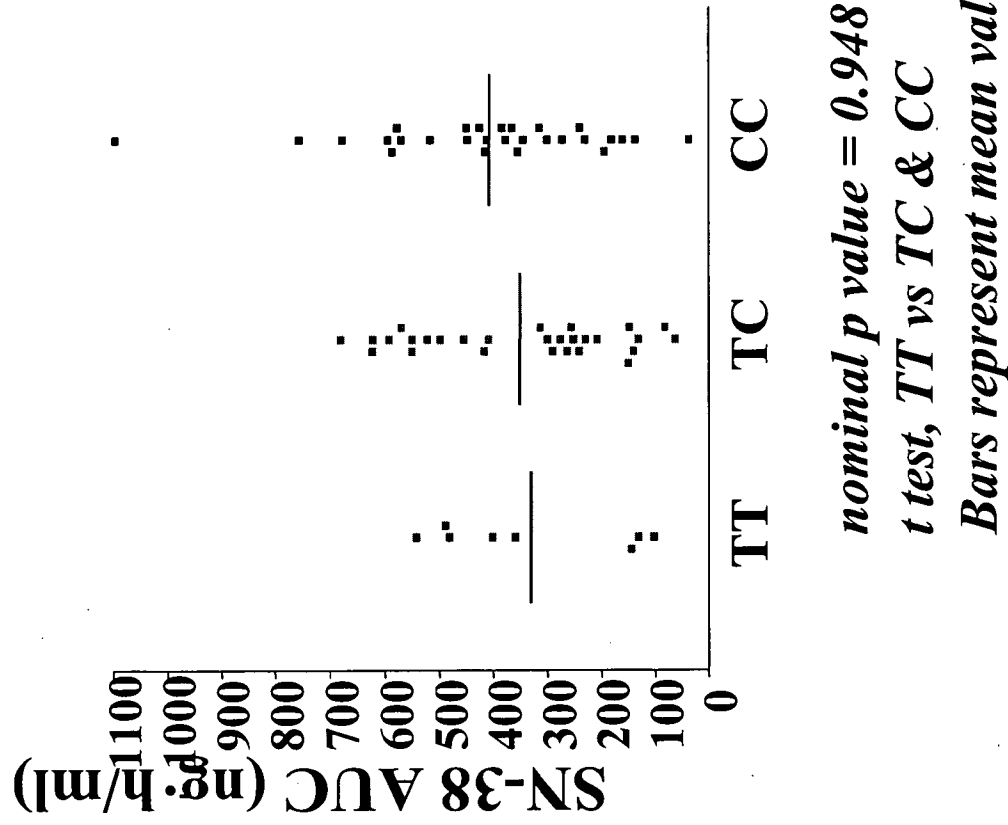


FIG. 2



030504

16369 U.S. PTO

PTO/SB/16 (08-03)

Approved for use through 07/31/2006. OMB 0651-0032

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EV 414834592 US

INVENTOR(S)					
Given Name (first and middle [if any])		Family Name or Surname		Residence (City and either State or Foreign Country)	
Mark J.		Ratain		Chicago, IL	
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto					
TITLE OF THE INVENTION (500 characters max)					
Methods and Compositions Relating to the Pharmacogenetics of ABCC2 Gene Variants					
Direct all correspondence to: CORRESPONDENCE ADDRESS					
<input checked="" type="checkbox"/> Customer Number: <div style="border: 1px solid black; padding: 5px; display: inline-block;">32425</div>					
OR					
<input type="checkbox"/> Firm or Individual Name					
Address					
Address					
City		State		Zip	
Country		Telephone		Fax	
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/> Specification Number of Pages <u>57 + 12-pg. sequence listing</u>		<input checked="" type="checkbox"/> CD(s), Number <u>1</u>			
<input checked="" type="checkbox"/> Drawing(s) Number of Sheets <u>2</u>		<input checked="" type="checkbox"/> Other (specify) <u>Statement re: sequence listing</u>			
<input type="checkbox"/> Application Date Sheet. See 37 CFR 1.76		(w/paper copy & disk); postcard			
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT					
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.		<div style="border: 1px solid black; padding: 10px; text-align: center;">FILING FEE Amount (\$) 80.00</div>			
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees.					
<input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: <u>50-1212</u>					
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.					
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input type="checkbox"/> No.					
<input checked="" type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: <u>National Institutes of Health,</u> <u>Grant Number GM61393</u>					

17548 U.S. PTO
60/550268

030504

[Page 1 of 2]

Respectfully submitted,

Date March 5, 2004

SIGNATURE

TYPED or PRINTED NAME Gina N. ShishimaTELEPHONE 512-536-3081REGISTRATION NO. 45,104

(if appropriate)

Docket Number: ARCD:405USP1**USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT**

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop Provisional Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Additional Page

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

ARCD:405USP1Chicago, IL

Number 2 of 2

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.